

REMARKS

Reconsideration is requested.

The specification has been amended to include the attached Sequence Listing and sequence identifiers corresponding to same. The attached paper and computer readable copies of the Sequence Listing are the same. No new matter has been added.

The Examiner interview of July 31, 2008 is acknowledged, with appreciation. The Interview Summary mailed September 30, 2008 is accurate in the brief description of the issues discussed during the interview.

Claims 1-4, 10-17, 19-23, 29-34 and 43-56 are pending. Claims 53-56 have been added. The new claims find support throughout the specification. No new matter has been added. Claims 53-56 are believed to read on the elected subject matter.

The claims have been revised to read on the elected subject matter. The applicants note, for the Examiner's convenience, that SEQ ID NO:5 is the motif QALGGH, SEQ ID NO:7 is an EAR motif, SEQ ID NO:8 is a B-Box motif and SEQ ID NO:9 is an L-Box motif. SEQ ID NO:1 encodes a sequence containing the QALGGH motif, an EAR motif, a B-Box motif and an L-Box motif, as described, for example, in Figure 3 of the specification.

The above amendments assume that the Amendment of March 31, 2008 has been entered. The Examiner is requested to advise the undersigned in the event anything further is required for entry of the above claim amendments.

The Examiner is requested to return a completely-initialed copy of the PTO-1449 Form bearing the OIPE date-stamped of June 7, 2005. Specifically, the PTO-1449

Ann Isabel SANZ MOLINERO
Appl. No. 10/537,897
Atty. Ref.: 4982-5
Amendment
November 18, 2008

Form returned with the Office Action of November 30, 2007, does not include the Examiner's initials next to foreign patent documents 96/39020 A, 1 230 843 A and 01/90343 A. The entirety of the PTO-1449 Form has been initialed by the Examiner on November 19, 2007, and the other references indicated as having been specifically considered by initialing next to each of the references. The above-noted foreign documents however do not include in the left-hand column the Examiner's initials. A completely-initialed copy of the PTO-1449 Form, pursuant to MPEP § 609, is requested.

The Examiner is requested to provide a complete PTO 892 Form, which includes the title of each cited Non-Patent Document so that the absence of same does not delay printing of the present application, or applications based on same, as a patent once patentable subject matter is identified and formal requirements are otherwise met. Moreover, lack of titles of each cited Non-Patent Document in an examiner's PTO-892 can, in the undersigned's experience, cause difficulty in prosecution of subsequent related applications where a different examiner may believe a title is required for a reference which is not required to be submitted in the subsequent application and may be difficult to locate. The Examiner is again requested to see MPEP § 707.05(e) in this regard.

The specification has been revised in the Amendment of March 31, 2008 to delete the browser executable code. Withdrawal of the objection to the specification is requested.

The objection to claim 4 is obviated by the above amendments filed March 31, 2008. Withdrawal of the objection is requested.

The Section 112, second paragraph, rejection of claims 10-17 is obviated by the amendments filed March 31, 2008. Withdrawal of the Section 112, second paragraph, rejection is requested.

To the extent not obviated by the above amendments, the Section 112, first paragraph "written description", rejection of claims 1-4, 10-17, 19-23 and 29-34 is traversed. Reconsideration and withdrawal of the rejection are requested in view of the above and the following comments.

The applicants provide, for example, on page 14 of the application the essential regions set forth in SEQ ID NOs: 5, 7, 8 and 9 identified in SEQ ID NO: 2, which is the protein encoded by SEQ ID NO: 1.¹ One of ordinary skill in the art will appreciate from the present specification that it is not necessary to include all the regions present in order to perform the methods of the claimed invention.

The Examiner asserts that the specification allegedly fails to adequately describe polynucleotide sequences that encode any 2XC2H2 zinc finger protein or a portion of SEQ ID NO: 1 having the same activity as SEQ ID NO: 1. Page 6, lines 14-23 of the application describes variants of 2XC2H2 nucleic acids and their encoded proteins useful in performing the methods of the claimed invention and therefore having a similar activity to SEQ ID NO: 1. Further the variants are described in detail on pages 6 to 16 of the present application. Page 7 details variant 2xC2H2 zinc finger nucleic acid having similar activity to SEQ ID NO: 1. Moreover, examples of SEQ ID NO: 2 homologous and

¹ "Particular zinc finger homologues useful in the methods of the present invention have one or more of the conserved motifs as depicted in SEQ ID NO 5, 6, 7, 8 and 9, or motifs that are 80% identical to these motifs or motifs that have conserved substitutions of amino acids. The 2xC2H2 protein as set

orthologous proteins, and the nucleic acids encoding the same are given in SEQ ID NO: 12 to SEQ ID NO: 25 which are further detailed in page 14 line 32-37 and page 15 lines 1-2. Examples of homologues of SEQ ID NO: 2 and the encoding polynucleotides originating from the same species are given in SEQ ID NO: 26 to SEQ ID NO: 35 and further detailed in page 15 line 6-17. Furthermore examples of alternative cDNAs (polynucleotides) encoding SEQ ID NO: 2 and therefore having the same activity of SEQ ID NO: 1, which also encodes SEQ ID NO: 2, are given in page 15 lines 23-24. The last paragraph of page 15 provides a number of 2xC2H2 homologous proteins useful in the methods of the invention. Identifying the polynucleotides encoding such homologous proteins is in the realm of the skill in the art.

In referring to Decisions of the Federal Circuit Court of Appeals, the Examiner states that a written description of an invention “requires a precise definition, one that defines structural features of the chemical genus that distinguishes it from other chemical structures”. See page 4 of the Office Action dated November 30, 2007. The Examiner further states that the description of a genus of cDNAs may be achieved by reciting a representative number of nucleotide sequences falling in the scope of the genus. See paragraph spanning pages 4-5 of the Office Action of November 30, 2007.

The applicants submit that the present specification provides the information which the Examiner asserts to be required.

Page 15 of the application, for example, provides several cDNAs encoding the same protein as that of SEQ ID NO 2. Numerous examples of 2XC2H2 encoding

forth in SEQ ID NO 2 comprises all the boxes as set forth in SEQ ID NO 5, 7, 8 and 9. All its paralogues

nucleotide sequences are described in the application, for example on page 15, five distinct examples of 2XC2H2 nucleotide sequences encoded in the genome of *Arabidopsis thaliana* are described. An exhaustive list is not believed to be required in the specification. The applicants believe however that the specification provides a sufficient representative number of examples that when taken together with those provided on page 15 and on Table 9 of the application, and the sequence listing, the figures, and the guidance provided on how to find other candidate sequences provides a precise definition of the 2XC2H2 genus, such that one of ordinary skill in the art will appreciate that the applicants were in possession of the claimed invention at the time the application was filed.

Beyond the above, the applicants note that page 14 of the application, for example, describes the structural features of SEQ ID NO: 2 as set forth in SEQ ID NO 5, 7, 8 and 9. Detailed characterization of each of these features and ways to identify them is provided in pages 11 to 13 of the application. Furthermore a substantial number of examples of 2XC2H2 nucleic acids and 2XC2H2 proteins comprising the structural features are cited in the patent application and sequence listing.

Further, attached as Annex 3 is a sequence alignment of multiple sequences of 2XC2H2 proteins as found in the sequence listing wherein positional motifs as represented by SEQ ID NOs: 5, 7, 8 and 9 are indicated by a box over the consensus sequence. Moreover, the attached Annex 4 provides a an alignment of multiple 2XC2H2 proteins as described in the specification in the paragraph bridging pages 15-

and orthologues also comprise all of these boxes.”

16 with reference to accession numbers known in the art. The position of the motifs as represented by SEQ ID NOs: 5, 7, 8 and 9 are indicated by a box over the consensus sequence. Annex 6 attached hereto provides a calculation for sequences of the disclosure containing the motifs related to SEQ ID NOs: 7, 8 and 9 of the specification and claims.

In summary, the applicants submit that there is ample written description provided in the present application providing ample structural information and a wealth of examples of polynucleotide sequences encoding 2XC2H2 proteins. Altogether, the structural features provided in the specification, the sequence in the sequence listing, and the guidance provided in the specification will lead one of ordinary skill in the art to appreciate that the applicants were in possession of the claimed invention at the time the application was filed

Withdrawal of the Section 112, first paragraph "written description", rejection is requested.

To the extent not obviated by the above amendments, the Section 112, first paragraph "enablement", rejection of claims 1-4, 10-17, 19-23 and 29-34 is traversed. Reconsideration and withdrawal of the rejection are requested in view of the above and the following comments.

The applicants submit that one of ordinary skill in the art will be able to make and use the claimed invention without undue experimentation. The specification provides an enabling disclosure of the claimed invention.

The Examiner is requested to see the attached Annex 1 in this regard which provides further experimental data showing that polynucleotide sequences encoding proteins having less than 100% identity to SEQ ID NO: 2, such as SEQ ID NO: 26 and SEQ ID NO: 36, when transformed into a plant give plants having increased yield, increased leaf surface and prolonged vegetative growth. The identity between SEQ ID NO: 2 and SEQ ID NO: 27 (which is encoded by the nucleic acid in SEQ ID NO: 26) is 42.2%. The identity between SEQ ID NO: 2 and SEQ ID NO: 37 (which is encoded by the nucleic acid in SEQ ID NO: 36) is 31 % (Annex2).

Despite the relatively low overall sequence identity, the proteins encoded by SEQ ID NO: 26 and SEQ ID NO: 36 comprised one or more of the regions identified in the claims.

The applicant has transformed rice plants with two different nucleic acids, one originating from Arabidopsis and a second one from rice, both encoding a 2XC2H2 protein. The sequences are described in the present application as SEQ ID NO: 26 and SEQ ID NO: 36, respectively. All experiments were carried out essentially as described in the Examples section of the application. The transgenic rice plants expressing the Arabidopsis 2XC2H2 transgene AtSTZparalog1 (SEQ ID NO: 26) showed increases in the parameters described in the Table I and Table II of Annex 1.

The transgenic rice plants expressing the rice 2XC2H2 transgene showed increases in the parameters described in the Table III of Annex 1.

Annex 5 attached hereto further describes transformation of rice plants with a nucleic acid as described in SEQ ID NO:28 of the present specification and which

encodes the polypeptide described in SEQ ID NO:29. The results of Annex 5 demonstrate an increased yield.

The Examiner is understood to be of the opinion that the state-of-the-art is such that one of skill in the art cannot predict which nucleic acid or portions thereof capable of hybridizing to SEQ ID NO:1 will encode a protein with the same activity as the protein encoded by SEQ ID NO: 1. The applicant respectfully disagrees and the Examiner is requested to see, for example, pages 13-16 of the present application which provides ample teaching of the amino acid sequence and structural features comprised within proteins encoded by a nucleic acid having similar activity to SEQ ID NO: 1 and therefore useful in the methods of the claimed invention.

Methods for the identification of 2xC2H2 zinc finger proteins having similar amino acid sequence and structural features to SEQ ID NO: 2 are well known in the art and provided in the present application. The Examiner is requested to see, for example, pages page 10 and page 11 first and second paragraphs in this regard. Furthermore, nucleic acid or portions thereof capable of hybridizing to SEQ ID NO:1 and encoding a protein with the same activity as the protein encoded by SEQ ID NO: 1 are illustrated by numerous examples in Page 7, second paragraph.

The Examiner is further understood to be of the opinion that methods for the prediction of protein structure from sequence data in order to ascertain functional aspects of the protein are extremely complex and that the prediction possibilities of the positions within the protein's sequence where amino acid substitutions can be made with a reasonable expectation of maintaining function are limited. Though there may be

examples in the field that can support such statements, the attention of the Examiner is brought to the fact that such general statements do not apply to the presently claimed invention.

Specifically, the methods for the prediction of protein structural features referred to in pages 10 and 11 of the application will not require undue experimentation and are straight forward to use. Such methods are routinely used in a successful manner by those in the art without the need for any non-routine work or undue burden. Performance of such methods is in the realm of those skilled in the art.

The applicants further understand the Examiner to be of the opinion that the application fails to provide guidance for selecting a sequence that gives results when transformed into a plant.

The applicant respectfully submits however that detailed guidance for searching and identifying sequences encompassed within the claims is given in the specification and the Example section illustrates/demonstrates the claimed invention in detail. As shown in attached Annex 1, the applicant has used the method disclosed in the application for finding homologous sequences to SEQ ID NO: 1 encoding 2XC2H2 proteins in Arabidopsis and rice genomes. Further the applicant has succeeded to isolate such homologous sequences (SEQ ID NO: 26 and SEQ ID NO: 36) and show that when introduced in a transgenic plant, the yield and the leaf surface are increased and the vegetative growth phase is prolonged.

Examples in Annex 1 further illustrate how the guidance provided in the present application is sufficient to enable the ordinarily skilled artisan to identify sequences useful in the methods of the claimed invention.

The Examiner is further understood to be of the opinion that transforming plants with heterologous genes involved in plant development produce unpredictable results. Though there maybe examples in the field that illustrate such a statement, this is not the rule.

A basic axiom in the genetics field is that which associates a phenotype to a genotype. The assumption that by transforming plants with a heterologous genes gives reproducible results is a basic principle widely used in the transgenic research field for example to prove the effect of a gene.

There are numerous examples to illustrate that transformation of a heterologous gene in a plant results in reproducible effects. For example the Arabidopsis gene AtNHX1 was reported to give salt-tolerance when transformed in Arabidopsis, such effect was reproduced when transformed in tomato and in *Brassica napus*. Further the transformation of homologous genes to AtNHX1 originating from *Atriplex gmelini* in rice plants resulted in the predicted effect of improving salt tolerance.

Another example is that of the Arabidopsis thaliana CBF1 transcription factor, which when transformed into Arabidopsis thaliana, Brassica napus and tomato plants consistently resulted in abiotic stress tolerance; and produced the same effect, that is, increased stress tolerance.

In summary, the specification as filed provides sufficient guidance to identify, isolate and use 2XC2H2 nucleic acids useful in the methods of the claimed invention, without undue experimentation. Furthermore the data in Annex 1 provide evidence that the predicted result of increasing expression of such nucleic acids in a plant is enabled over the scope of the claims.

On page 10 of the Office Action dated November 30, 2007, the Examiner argues lack of disclosure of methods to isolate sequences encompassed by the claims. The Examiner is requested to see however the present specification where methods to search and identify the sequences encompassed by the claims are described, as detailed above. Several alternative methods are set forth. Several examples are provided. Alternative methods to isolate nucleic acids of interest are well known in the art (Sambrook et al 2001, page 8, line 6 of the Application). Further and as shown in Annex 1, the applicant has used the method disclosed in the application for identifying homologous sequences to SEQ ID NO: 1 encoding 2XC2H2 proteins in Arabidopsis and rice genomes. Furthermore the applicant has isolated such homologous sequences (SEQ ID NO: 26 and SEQ ID NO: 36) and shown that when introduced in a transgenic plant, the yield and the leaf surface are increased and the vegetative growth phase is prolonged. Examples in Annex 1 further illustrate how the guidance provided in the application is sufficient to enable one of ordinary skill in the art to identify sequences useful in the methods of the claimed invention. Altogether, the teaching and guidance in the specification and the illustration with the various examples fully enables the

Ann Isabel SANZ MOLINERO
Appl. No. 10/537,897
Atty. Ref.: 4982-5
Amendment
November 18, 2008

ordinarily skilled artisan to isolate and identify the sequences encompassed in the claims.

Withdrawal of the Section 112, first paragraph "enablement", rejection is requested.

The Section 102 rejection of claims 1-4, 10-17, 19-23 and 29-34 over Pineda (WO 01/36598 A1), is obviated by the above amendments. Reconsideration and withdrawal of the rejection are requested as the cited document fails to teach or suggest, for example, selecting for plants having increased yield, as required by the claims.

The claims are submitted to be in condition for allowance and a Notice to that effect is requested.

The Examiner is requested to contact the undersigned, preferably by telephone, in the event anything further is required in this regard.

Respectfully submitted,

NIXON & VANDERHYE P.C.

By: /B. J. Sadoff/
 B. J. Sadoff
 Reg. No. 36,663

BJS:
901 North Glebe Road, 11th Floor
Arlington, VA 22203-1808
Telephone: (703) 816-4000
Facsimile: (703) 816-4100

Annex 1

Example A

AtSTZparalog1: SEQ ID NO: 26 under the control of constitutive promoter GOS2

A DNA fragment encoding a 2XC2H2 protein represented in the application as filed by SEQ ID NO: 27 was isolated from an *Arabidopsis thaliana* seedling cDNA library (Invitrogen, Paisley, UK) by PCR amplification and subsequent cloning in an entry clone vector according to the methods described in Example 1 of the present application. SEQ ID NO: 27 polypeptide was encoded by the longest open reading frame of SEQ ID NO: 26 (AtSTZparalog1).

The primers used for the PCR amplification were as follows:

Forward primer: Ggggacaagttgtacaaaaaagcaggcttaacaatggccctcgaagcg

Reverse primer: Ggggaccacttgtacaagaaagctgggttcgagtattagatttttaagataaatc

The entry clone was subsequently used in an LR reaction with a destination vector used for rice transformation to generate the plant expression vector pGOS2::AtSTZparalog1. The constitutive promoter, GOS2, was mentioned in Table 10 on page 48 of the application as filed as being a promoter useful in the methods of the invention.

Phenotypic characterization of the transformed plants was carried out essentially as described in Example 3 of the present application. The results are shown in Table I below.

Table I: Results of phenotypic characterization of T2 rice plants transformed with pGOS2::AtSTZparalog1.

| pGOS2::AtSTZparalog1 | |
|----------------------|-----------------------------------------------------------------------|
| Parameter | % increase in the transgenic plants compared to the nullzygous plants |
| Aboveground area | 10 |
| Root Area | 4 |

| pGOS2::AtSTZparalog1 | |
|------------------------|-----------------------------------------------------------------------|
| Parameter | % increase in the transgenic plants compared to the nullzygous plants |
| Total Seed Weight | 48 |
| Number of filled seeds | 46 |
| Total number of seeds | 12 |
| Seed filling rate | 30 |
| Flowers per panicle | 8 |
| Harvest index | 35 |
| Days to flowering | 4 |

The above results show that overexpression of the nucleic acid represented by SEQ ID NO: 27 (encoding the 2XC2H2 zinc finger protein represented by SEQ ID NO 26) under the control of a constitutive promoter (GOS2) gives:

- Increased plant yield (in the form of increased root area, total seed weight, total number of seeds, number of filled seeds, seed filling rate, flowers per panicle and harvest index);
- Increased leaf surface area (as manifested by increased aboveground area); and
- Prolonged vegetative growth (where the time to flower was on average 4% longer in transgenic plants compared to corresponding nullzygous plants).

Example B

AtSTZparalog1: SEQ ID NO: 26 under the control of seed-specific promoter prolamin

A DNA fragment encoding a 2XC2H2 protein represented by SEQ ID NO: 27 was isolated from an *Arabidopsis thaliana* seedling cDNA library (Invitrogen, Paisley, UK) by PCR amplification and subsequent cloning in an entry clone vector according to the methods described in Example

1 of the present application. SEQ ID NO: 27 polypeptide was encoded by the longest open reading frame of SEQ ID NO: 26 (AtSTZparalog1).

The primers used for the PCR amplification were as follows:

Forward primer: Ggggacaagtttgtaaaaaagcaggcttaacaatggccctcgaagcg

Reverse primer: Ggggaccactttgtacaagaaagctgggttcgagtattagatttttaagataaatc

The entry clone was subsequently used in an LR reaction with a destination vectors used for rice transformation to generate the plant expression vector pPROLAMIN::AtSTZparalog1. The seed-specific promoter, prolamin, was mentioned in Table 10 on page 48 of the application as filed as being a promoter useful in the methods of the invention.

Phenotypic characterization of the transformed plants was carried out essentially as described in Example 3 of the present application. The results are shown in Table II below.

Table II: Results of phenotypic characterization of T2 rice plants transformed with pPROLAMIN::AtSTZparalog1.

| pPROLAMIN::AtSTZparalog1 | |
|--------------------------|------------------------------------------------------------------------|
| Parameter | % increase in the transgenic plants compared to the nullyzygous plants |
| Aboveground area | 4 |
| Total Seed Weight | 8 |
| Number of filled seeds | 7 |
| Total number of seeds | 5 |
| Flowers per panicle | 4 |
| Harvest index | 5 |

The above results show that overexpression of the nucleic acid represented by SEQ ID NO: 27 (encoding the 2XC2H2 zinc finger protein represented by SEQ ID NO 26) under the control of a seed-specific promoter (prolamin) gives:

- Increased plant yield (in the form of increased aboveground area, total seed weight, total number of seeds, number of filled seeds, flowers per panicle and harvest index);
- Increased leaf surface area (as manifested by increased aboveground area); and

Example C

OsSTZ(ortholog): SEQ ID NO: 36 under the control of root-specific promoter RCc3

A DNA fragment comprising the coding region of SEQ ID NO 36 was PCR amplified and cloned using methods essentially as described in the Examples section of the present application.

The primers used for the PCR amplification were as follows:

Forward primer: ggggacaagttgtacaaaaagcaggctaaacaatgtcgagcgctcgt

Reverse primer: ggggaccactttgtacaagaaagctgggtctgaattacgcggtgagaag

A plant transformation vector carrying the coding region of SEQ ID NO 36 under the control of the root specific promoter, RCc3, was made, giving construct CD10315 described in Table 9 on page 47 of the present application. Agrobacterium-mediated transformation of rice plants was carried out to generate transgenic rice plants carrying the construct CD10315.

The results of the phenotypic evaluation of the CD10315-transgenic plants of the T2 generation are shown in Table III below.

Table III: Results of phenotypic characterization of T2 rice plants transformed with pRCc3::OsSTZortholog.

| CD10315 plants | |
|------------------------|-----------------------------------------------------------------------|
| Parameter | % increase in the transgenic plants compared to the nullzygous plants |
| Aboveground area | 6 |
| Root Area | 5 |
| Total Seed Weight | 29 |
| Number of filled seeds | 29 |
| Seed filling rate | 16 |
| Total number of seeds | 17 |
| Harvest index | 25 |

The above results show that overexpression of a nucleic acid encoding the 2XC2H2 zinc finger protein represented by SEQ ID NO 36) under the control of a root-specific promoter (RCc3) gives:

- Increased plant yield (in the form of increased aboveground area, root area, total seed weight, total number of seeds, number of filled seeds, seed filling rate and harvest index); and
- Increased leaf surface area (as manifested by increased aboveground area).

Annex 2

Sequence identity between SEQ ID NO: 2 and SEQ ID NO: 27 and SEQ ID NO: 37.

Amino acid sequence identity between SEQ ID NO: 2 and SEQ ID NO: 27 (encoded protein in SEQ ID NO: 26) and SEQ ID NO: 37 (encoded protein in SEQ ID NO: 36) was determined using the Needleman and Wunsch as described in page 10 paragraph 2 of the Application.

Results for the comparison between SEQ ID NO: 2 and SEQ ID NO: 27 are given in Table 1 and Figure 1.

Results for the comparison between SEQ ID NO: 2 and SEQ ID NO: 37 are given in Table 2 and Figure 2.

Sequence identity between SEQ ID NO: 2 and its homologous protein SEQ ID NO: 27 is 42.2 %. Highly conserved domains corresponding to the SEQ ID NO 5 (motif within C2H2 zinc finger domain), 7 (EAR motif), 8 (B-box) and 9 (L-box) are readily identifiable in both SEQ ID NO: 27 and 37.

Table 1. Similarity and identity between SEQ ID NO: 2 and SEQ ID NO: 27.

| Algorithm | Query 1 | Query 1 | % identity | % gaps | % similarity |
|----------------------|-----------|----------|------------|--------|--------------|
| Needleman and Wunsch | SEQIDNO27 | SEQIDNO2 | 42.2 | 22 | 53.5 |

| | | | |
|-----------|-----|-----------------------------------------------------|-----|
| SEQIDNO27 | 1 | MALEAMNTPTSSFTRIETKEDLMNDAVF--IEPWLKRKRKRQRSHSPSS | 48 |
| | | :..: ...:..... : : | |
| SEQIDNO2 | 1 | MALEALTSPRLASPIPLFED---SSVFHGVHWTGKRKRKRSR----- | 42 |
| SEQIDNO27 | 49 | SSSSPPRSRPKSQNQDLTEEEYLALCLLMLAKD--QPSQTRFHQQSLSLT | 96 |
| | | ...: : . : : : | |
| SEQIDNO2 | 43 | -----DFHHQNLTEEEYLAFCLMLLARDNRQP-----P | 70 |
| SEQIDNO27 | 97 | PPPEKNLPYKCNVCEKAFPSYQALGGHKASHRIKPPTVISTTADD---S | 143 |
| | | : :: ... | |
| SEQIDNO2 | 71 | PPPAVEKLSYKCSVCDKTFSSYQALGGHKASHRKNLSQTLGGGDDHSTS | 120 |
| SEQIDNO27 | 144 | TAPTISIVAGEKHPIAASGKIHECSICHKVFP TGQALGGHKRCHYEGNLG | 193 |
| | | : : : . : : : : : : : : : : : | |

| | | | |
|-----------|-----|-----------------------------------------------------------------------------------------|-----|
| SEQIDNO2 | 121 | SATTTSAVT-----TSGSKSHVCTICNKSFPSPGQALGGHKRCHYEGR-- | 162 |
| SEQIDNO27 | 194 | GGGGGGSKSISHSGSVSTVSEERSHRGFIDLNLPAPELSLHHNPVIDE | 243 |
| | | : : : : . : : : . : : | |
| SEQIDNO2 | 163 | --NNINTSSVSNSEGAGSTSHVSSSHRGF-DLNIPPIPEFSMVNG---DD | 206 |
| SEQIDNO27 | 244 | EILSPLTGKKPLLLTDHDQVIKEDLSLKI-- | 273 |
| | | : : : . . . : . . : : | |
| SEQIDNO2 | 207 | EVMSPMPAKKP-----RFDFPVKLQL | 227 |

Figure 1. Alignment between SEQ ID NO: 2 and SEQ ID NO: 27.

Table 2. Similarity and identity between SEQ ID NO: 2 and SEQ ID NO: 27.

| Algorithm | Query 1 | Query 1 | % identity | % gaps | % similarity |
|----------------------|-----------|----------|---------------|--------|-----------------|
| Needleman and Wunsch | SEQIDNO27 | SEQIDNO2 | 31 | 31 | 42.2 |

| | | | |
|-----------|-----|-----------------------------------------------------|-----|
| SEQIDNO37 | 1 | MSSASSMEALHAAVLKEEQQHEVEEATVVTSSSATSGEEGGHLF---- | 45 |
| | | :...: | |
| SEQIDNO2 | 1 | -----MALEALTSPLASPIPLFED | 21 |
| SEQIDNO37 | 46 | -----QGWAQRKRSRQRS-----EEENLALCLMLLARGGHHRVQ | 80 |
| | | :. . . : . . . :: ... | |
| SEQIDNO2 | 22 | SSVFHGVHWTGKRSKRSRSDFHQNLTEEEYLAFCLMLLARDNR---Q | 68 |
| SEQIDNO37 | 81 | APPPLSASAPPAGAEFKCSVCCKSFSSYQALGGHKTSHRVKLPTPPAAP | 130 |
| | |: . : | |
| SEQIDNO2 | 69 | PPPP-----PAVEKLSYKCSVCCKTFSSYQALGGHKASHRKNLS----- | 107 |
| SEQIDNO37 | 131 | VLAPAPVAALLPSAEDREPATSTTAASSDGMTNRVHRCSICQKEFPTGQA | 180 |
| | |:.....: .: .: .: .: .: | |
| SEQIDNO2 | 108 | -----QTLSGGGDDHSTSSATTTSAVTTGSGKSHVCTCNKSFPSGQA | 150 |
| SEQIDNO37 | 181 | LGGHKKRKHVDGGVGAGAGASSTELLATVAASEVVGSSGNGQSATRAFDLN | 230 |
| | | . : : :..... :..... . | |
| SEQIDNO2 | 151 | LGGHKKRCHYEGNNNINT-----SSVSNSEGAGSTSHVSSSHRGFDLN | 192 |
| SEQIDNO37 | 231 | LPAPVEFVWRPCSKGKMWDEEEVQSPLAFKKPRLTLA----- | 269 |
| | | : : : .:. : : : : : : : : : : | |
| SEQIDNO2 | 193 | IPPIPEF-----SMVNGDDEVMSMPAKKPRDFPVKLQL | 227 |

Figure 2. Alignment between SEQ ID NO: 2 and SEQ ID NO: 37.

Annex 3

Multiple amino acid sequence alignment of 2XC2H2 proteins as found in the in the amino acid of the sequence listing of the application as filed. The position of the motifs as represented by SEQ ID NO, 5, 7, 8 and 9 are indicated by a box over the consensus sequence.

```

1                                     50
SEQIDNO_11_Datisca_glomerata (1) ---MALEALNSPTT-----ATPVF--H---YDDP-SLNVLEPWTK
SEQIDNO_13_Glycine_max (1) ---MALEALNSPTT-----TAPSF--P-----PDDFTIPWAK
SEQIDNO_15_Medicago_sativa (1) ---MALEALNSPTT-----ATPFT--P---FEEP-N-LSYLETPWTK
SEQIDNO_17_Nicotiana_tabacum (1) ---MTLEALKSPTA-----ATPTL--PPRYEDDD-EIHNDSWAK
SEQIDNO_19_Oryza_sativa (1) MSSASSMEALHAAVLKEEQQQHEVEEATVVTSSSATSGEEGCHLPQGWAK
SEQIDNO_37_Oryza_sativa (1) MSSASSMEALHAAVLKEEQQQHEVEEATVVTSSSATSGEEGCHLPQGWAK
SEQIDNO_23_Triticum_aestivum (1) -MSSSAMREALHALIP---EQHQLDVEAAA-VSSATSGEESGHVLCQWAK
SEQIDNO_27_Arabidopsis_thaliana (1) ---MALEAMNTPTS-----SFTRI--ETKEDLMN-DAVFTEPWLK
SEQIDNO_31_Arabidopsis_thaliana (1) ---MALETLNSPTS-----TTTTTAPPPFLRCLDETEFENESWTK
SEQIDNO_35_Arabidopsis_thaliana (1) ---MALETLNSPTA-----TTTAR--PLLRYREEMEFENLEQWAK
SEQIDNO_39_Arabidopsis_thaliana (1) -----
SEQIDNO_50_Arabidopsis_thaliana (1) ---MGQDFVCGSDQTQIIKGRKTKRQRSS--STFVVTAATVTSTSSSAG
SEQIDNO_46_Arabidopsis_thaliana (1) -----
SEQIDNO_21_Petunia_hybrida (1) ---MALEALNSPT-----TTTTP--SFQFENNG--LKYLESWTK
SEQIDNO_25_Capsicum_annuum (1) ---MALEALNSPTG-----TPTTP--PFQFESDGGQLRYLENWRK
SEQIDNO_29_Arabidopsis_thaliana (1) ---MALETLTSPRL-----SSPMP--T-----LPQDSALCEHG
SEQIDNO_33_Arabidopsis_thaliana (1) ---MALEALNSPRL-----VEDP-----LRFNGVEQWTK
Consensus (1) MALEALNSPT L W K

51                                     100
SEQIDNO_11_Datisca_glomerata (32) RK--RSKRTRLDSP-----HTEBEYLALCLIMLARGR
SEQIDNO_13_Glycine_max (28) RK--RSKRSRDHPS-----TEBEYLALCLIMLARGG
SEQIDNO_15_Medicago_sativa (33) GK--RSKRSRMDQSS-----CTEBEYLALCLIMLARSG
SEQIDNO_17_Nicotiana_tabacum (35) GK--RSKPRIDAP-----PTEBEYLALCLIMLARSG
SEQIDNO_19_Oryza_sativa (51) RK--RSKRQRS-----EBENLALCLIMLARSG
SEQIDNO_37_Oryza_sativa (51) RK--RSKRQRS-----EBENLALCLIMLARGG
SEQIDNO_23_Triticum_aestivum (46) RK--RSKRQRS-----EBENLALCLIMLARSG
SEQIDNO_27_Arabidopsis_thaliana (35) RK--RSKRQRSHSPSSSSSSPPRSRPRKSONQDIETEBEYLALCLIMLARQD
SEQIDNO_31_Arabidopsis_thaliana (39) RK--RTKRHRIDQPNPP-----PSEBEYLALCLIMLARGS
SEQIDNO_35_Arabidopsis_thaliana (36) RK--RTKRQRFDHGHQNG-----ETNKNLPSBEBEYLALCLIMLARGS
SEQIDNO_39_Arabidopsis_thaliana (1) ----MKRDRSDYEES-----MKHIDI VESLIMLDRSF
SEQIDNO_50_Arabidopsis_thaliana (45) GS--GGERAVSDDEYN SAVSSPVITDCTQ-----EEEDMALCLIMLARGT
SEQIDNO_46_Arabidopsis_thaliana (1) ----MVARSEE-----VEIVEDTAAKCIMLSRGV
SEQIDNO_21_Petunia_hybrida (33) GK--RSKRQRSMERQ-----CTEBEYLALCLIMLARSD
SEQIDNO_25_Capsicum_annuum (36) GK--RSKRQRSMERQ-----PTEBEYLALCLIMLARSG
SEQIDNO_29_Arabidopsis_thaliana (29) SKGKRSKRQRSEFDR-----QSLTEBEYLALCLIMLARDC
SEQIDNO_33_Arabidopsis_thaliana (27) CK--RSKRQRSDLHHN-----HRLTEBEYLALCLIMLARDC
Consensus (51) K--RSKR RSD TEBEYLALCLIMLAR G
```

SEQ ID NO 8

SEQ ID NO 9

```

101                                     150
SEQIDNO_11_Datisca_glomerata (62) VA-----SANRRDSQSSIQIQP-----EATTSATKVSX
SEQIDNO_13_Glycine_max (56) TT-----TVNNRHVSPPLQEQPQ-----PTPDPSTKLSY
SEQIDNO_15_Medicago_sativa (64) NN-----NDKKSDSVATP-----LTTVKLSH
SEQIDNO_17_Nicotiana_tabacum (65) TG-----TRTGLTDATTSQQPADKKTAELPPVHKKEVATEQAQBSY
SEQIDNO_19_Oryza_sativa (76) HH-----RVQAPPPLSASAPPAG-----AEF
SEQIDNO_37_Oryza_sativa (76) HH-----RVQAPPPLSASAPPAG-----AEF
SEQIDNO_23_Triticum_aestivum (71) KQ-----RVQAPQPESFAAPVPA-----EF
SEQIDNO_27_Arabidopsis_thaliana (83) PS-----QTRFHQQSQSLTPEP-----ESKNLPY
SEQIDNO_31_Arabidopsis_thaliana (72) SD-----HMSPPSDHHSLSPLSDHQK-----DY
SEQIDNO_35_Arabidopsis_thaliana (76) AV-----QSPPLPPLPSRASPSDHR-----DY
SEQIDNO_39_Arabidopsis_thaliana (29) VVK-----QIDVKQSTGSKTNHNN-----HE
SEQIDNO_50_Arabidopsis_thaliana (87) VLPSPDLK--NSRKIHQKISSENS-----FYVY
SEQIDNO_46_Arabidopsis_thaliana (27) ECG-----GGGEKR-----VF
SEQIDNO_21_Petunia_hybrida (64) GS-----VNNRSRLPPPLPPSVPVTSQ-----INATLLE-QKNLY
SEQIDNO_25_Capsicum_annuum (67) GS-----VNHQRSLPPAPVMKLHAPSS-----SSAAEEKEKMYV
SEQIDNO_29_Arabidopsis_thaliana (64) DR-----NRDLPLPSSSSSEP-----LLPPLPTPIY
SEQIDNO_33_Arabidopsis_thaliana (62) G-----DLDSVTVAEKE-----SY
Consensus (101) P Y
```

| | | | |
|---------------------------------|-------|------------------------------------------------------|-----|
| SEQIDNO_11_Datisca_glomerata | (90) | KCSVCDKAFSSYQALGGHKASHRKLAGGED-----QSTSF---AT--TN | 200 |
| SEQIDNO_13_Glycine_max | (86) | KCSVCDKSFPSYQALGGHKASHRKLAGAAE-----DQPPS---T--TT | |
| SEQIDNO_15_Medicago_sativa | (85) | KCSVCDKAFSSYQALGGHKASHRKAVMSAT-----TAEDQ---IT--TT | |
| SEQIDNO_17_Nicotiana_tabacum | (106) | KCSVCDKAFSSYQALGGHKASHRKTTTTAT-----AASDDNNPST--ST | |
| SEQIDNO_19_Oryza_sativa | (98) | KCSVCGKSFPSYQALGGHKASHRKVLPTTPP-----AAPVLAPAPVAA--LL | |
| SEQIDNO_37_Oryza_sativa | (98) | KCSVCGKSFPSYQALGGHKASHRKVLPTTPP-----AAPVLAPAPVAA--LL | |
| SEQIDNO_23_Triticum_aestivum | (91) | KCSVCGKSFPSYQALGGHKASHRKVLPTTPP-----AAPVLAPAPVAA--LL | |
| SEQIDNO_27_Arabidopsis_thaliana | (107) | KCMVCEKAFPSYQALGGHKASHRIKPPPTVI-----STTADDS-TA--PT | |
| SEQIDNO_31_Arabidopsis_thaliana | (95) | KCSVCGKSFPSYQALGGHKASHRKPVSVDV-----NNGNGTVTNN--GN | |
| SEQIDNO_35_Arabidopsis_thaliana | (98) | KCTVCGKSFPSYQALGGHKASHRKPTNTSI-----TSGNQELSN--SH | |
| SEQIDNO_39_Arabidopsis_thaliana | (50) | ECKTCNRKFDSTQALGGHKASHRKPKLIVD----- | |
| SEQIDNO_50_Arabidopsis_thaliana | (114) | ECKTCNRFTFSSQALGGHKASHRKPRSTTEEKTRLPLTPKSSASEEGQN | |
| SEQIDNO_46_Arabidopsis_thaliana | (38) | ECKTCLEKFSSTQALGGHKASHRKKLINSDD-----P-----S- | |
| SEQIDNO_21_Petunia_hybrida | (99) | KCSVCGKGFPSYQALGGHKASHRKLVSMGG-----DEQSTTSTTT--NV | |
| SEQIDNO_25_Capsicum_annuum | (103) | KCSVCGKGFPSYQALGGHKASHRKLVPGG-----DDQSTTSTTT--NA | |
| SEQIDNO_29_Arabidopsis_thaliana | (90) | KCSVCDKAFSSYQALGGHKASHRKSFSLTQ-----SAGGD-ELST--SS | |
| SEQIDNO_33_Arabidopsis_thaliana | (76) | KCGVCYKTFSSYQALGGHKASHRSLYGGGE----- | |
| Consensus | (151) | KCSVC KSFSSYQALGGHKASHRK T | |

SEQ ID NO 5

| | | | |
|---------------------------------|-------|----------------------------------------------------|-----|
| SEQIDNO_11_Datisca_glomerata | (129) | SATV-----TTTASGG-GGRSHECSICHKSFPTGQALGGHKRCHYEG | 250 |
| SEQIDNO_13_Glycine_max | (124) | SSAA-----ATSSASG--GKAHECSICHKSFPTGQALGGHKRCHYEG | |
| SEQIDNO_15_Medicago_sativa | (124) | SSAA-----TTSSASNG-KNITHECSICHKSFPTGQALGGHKRCHYEG | |
| SEQIDNO_17_Nicotiana_tabacum | (148) | STGA-----VNISALNP-TGRSHVCSICHKAFPTGQALGGHKRCHYEG | |
| SEQIDNO_19_Oryza_sativa | (142) | PSAEDREPATSTAAASSDGMNVRHRCISICQKEFPTGQALGGHKRCHYEG | |
| SEQIDNO_37_Oryza_sativa | (142) | PSAEDREPATSTAAASSDGMNVRHRCISICQKEFPTGQALGGHKRCHYEG | |
| SEQIDNO_23_Triticum_aestivum | (139) | PSAEP---ATSTAAASSDGMNVRHRCISICQKEFPTGQALGGHKRCHYEG | |
| SEQIDNO_27_Arabidopsis_thaliana | (148) | ISIV-----ACEKHPIAASCKIHECSICHKVFPPTGQALGGHKRCHYEG | |
| SEQIDNO_31_Arabidopsis_thaliana | (137) | ISNG-----LVGQSGKTHNCSICPKSFPSGQALGGHKRCHYEG | |
| SEQIDNO_35_Arabidopsis_thaliana | (140) | SNSGSVVINVTVNTGNGVSCSKIHCTCSICPKSFASGQALGGHKRCHYEG | |
| SEQIDNO_39_Arabidopsis_thaliana | (80) | -Q-----PQVKHRNKENDMHKCTICDCMFGTGQALGGHKRCHYEG | |
| SEQIDNO_50_Arabidopsis_thaliana | (164) | SHFKVGSALASQASNIINKANKVHECSICGSEFTSGQALGGHMRRHRTA | |
| SEQIDNO_46_Arabidopsis_thaliana | (70) | LLG-----SLNKKTKTATSHPCPICGVEFPMGQALGGHMRRHRTSE | |
| SEQIDNO_21_Petunia_hybrida | (141) | TG-T-----SSANVNGN--GRTHECSICHKCFPTGQALGGHKRCHYEG | |
| SEQIDNO_25_Capsicum_annuum | (144) | TGTT-----TSVNGNGNRSGRTHECSICHKCFPTGQALGGHKRCHYEG | |
| SEQIDNO_29_Arabidopsis_thaliana | (131) | AITT-----SGISGGGGGSVKSHVCSICHKSFATGQALGGHKRCHYEG | |
| SEQIDNO_33_Arabidopsis_thaliana | (106) | -----NDKSTP-STAVKSHVCSVCGKSFATGQALGGHKRCHYEG | |
| Consensus | (201) | ST S SGKSH CSIC K FPTGQALGGHKRCHYDG | |

SEQ ID NO 5

| | | | |
|---------------------------------|-------|---------------------------------------------------|-----|
| SEQIDNO_11_Datisca_glomerata | (171) | SICGNSIHNNNTTNSGSGGMSMTSEVGSSTHTVSHSHRDFDLNIPALPE | 300 |
| SEQIDNO_13_Glycine_max | (164) | NGNGN-----NNNSNSVTVASEGVGSSTHTVSHGHHRDFDLNIPALPE | |
| SEQIDNO_15_Medicago_sativa | (166) | SVG-----AGAGAGSNAVTASEGVGLSHSHHRDFDLNIPALPE | |
| SEQIDNO_17_Nicotiana_tabacum | (190) | KLGGN---SRDLGGGGGGHSGSVLTTSDCGASTHTLDFDLNIPALPE | |
| SEQIDNO_19_Oryza_sativa | (192) | GVCAGAGASST-----ELLATVAASEVGSNGGQSATRAFDLNPAPPE | |
| SEQIDNO_37_Oryza_sativa | (192) | GVCAGAGASST-----ELLATVAASEVGSNGGQSATRAFDLNPAPPE | |
| SEQIDNO_23_Triticum_aestivum | (186) | --GVGAAASST-----ELLAAAAASEVGSSTGNGSSAARFDLNPAPPE | |
| SEQIDNO_27_Arabidopsis_thaliana | (191) | NLGGC-----GGGGSKSISHSGSVSSTVSEERSHRGFIDLNIPALPE | |
| SEQIDNO_31_Arabidopsis_thaliana | (175) | GN-----GNNGDNCHKFDLNLPAQVSDETIGK | |
| SEQIDNO_35_Arabidopsis_thaliana | (190) | GNN-----GNGNGSSNSVELVAGSDVSDMDNERW | |
| SEQIDNO_39_Arabidopsis_thaliana | (119) | MITEQS-----IVPSVVYSRPFVFNRCSSSKELLIDLNIPALPE | |
| SEQIDNO_46_Arabidopsis_thaliana | (214) | VTTISPVAATAEVS--RNSTEEIEINIGRSMEOQRKYLPLDLNIPAPGD | |
| SEQIDNO_21_Petunia_hybrida | (111) | KASPG-----TLVTRSELPETITVTLKKSSSGKRVACLDLDMES | |
| SEQIDNO_25_Capsicum_annuum | (181) | GNNGNG-----SVSVGVTSSEGVGSTISHHRDFDLNIPALPE | |
| SEQIDNO_29_Arabidopsis_thaliana | (187) | GIGNGNAN-----SGVGSASVGVTSSEGVGSTVS--HRDFDLNIPALPE | |
| SEQIDNO_33_Arabidopsis_thaliana | (174) | KNCGG-----VSSSVSNSEVVGSTSHVSSG--HRDFDLNIPALPE | |
| Consensus | (251) | G S S S SS S R PDLNIPALPE | |

SEQ ID NO 7

| | 301 | | 350 |
|---------------------------------|------------------------------------------------------|--|----------------|
| SEQIDNO_11_Datisca_glomerata | (221) ERSNFFISG-----DDEVESPHPAKKPRILMK----- | | |
| SEQIDNO_13_Glycine_max | (207) ESTKVGEDF-----VESPHPMKKPRIFVVPKIEIPQFQ---- | | |
| SEQIDNO_15_Medicago_sativa | (204) ESKKFFVDD-----EVFSPLPAKKKPCIEFKLEIPSHY----- | | |
| SEQIDNO_17_Nicotiana_tabacum | (236) LQLGLSIDCGRKSQQLPMVQEVESPMPAKKPRLLFSLG----- | | |
| SEQIDNO_19_Oryza_sativa | (237) EVWRPCSKG---KKMWDEEEVQSPLAPKKPRLLTA----- | | |
| SEQIDNO_37_Oryza_sativa | (237) EVWRPCSKG---KKMWDEEEVQSPLAPKKPRLLTA----- | | |
| SEQIDNO_23_Triticum_aestivum | (229) EVWRPCAKG---KMMWEDDEEVQSPLAPKKPRLLTA----- | | |
| SEQIDNO_27_Arabidopsis_thaliana | (233) LSLHHN-----PIVDEEELSDLTGKKPLLLTDHDQVIKKEDLSLKI | | |
| SEQIDNO_31_Arabidopsis_thaliana | (204) SQ-----LSGEETKSVL----- | | |
| SEQIDNO_35_Arabidopsis_thaliana | (220) SEE-----SAIGGHRGFDLNLDPADQVSVTTS----- | | |
| SEQIDNO_39_Arabidopsis_thaliana | (156) DLVLIFGKN-----LVPQIDLKFPVN----- | | |
| SEQIDNO_50_Arabidopsis_thaliana | (262) DLR-----ESKFGGVFSATPAEIDCHY----- | | |
| SEQIDNO_46_Arabidopsis_thaliana | (152) LVN-----WKIELGRTIS----- | | |
| SEQIDNO_21_Petunia_hybrida | (219) EWPGFGSGE-----DEVESPHPAKKSRGLPCKLELFKGL---- | | |
| SEQIDNO_25_Capsicum_annuum | (228) EWLGFSGSGE-----DEVESPHPAKKSRGLPCKYELFQH---- | | |
| SEQIDNO_29_Arabidopsis_thaliana | (212) ESMVNG-----DEVMSPMAPKKLRFDPEKP----- | | |
| SEQIDNO_33_Arabidopsis_thaliana | (175) ESP-----DEVMSPMATKKPREK----- | | |
| Consensus | (301) F | | EEV SPL KK RLL |

Annex 4

Multiple amino acid sequence alignment of 2XC2H2 proteins as found in the polypeptides recited by their sequence database accession number in lines 17 to 30 of page 15 and the paragraph bridging page 15 and 16 of the specification. The position of the motifs as represented by SEQ ID NO, 5, 7, 8 and 9 are indicated by a box over the consensus sequence.

With the aim to improve visualization of the Alignment, the sequences of BAA19926, BAA20137, BAA21919, NP_182037, BAA19114, CAA18741 and CAB80245 were truncated at their N-terminus.

```

                                301                                350
AAK01713      (1)  -----MSSASSMEALHAAVLKEEQQQHEVEEATVVTSSS
AF332876      (1)  -----MSSASSMEALHAAVLKEEQQQHEVEEATVVTSSS
AAM67193      (1)  -----MALEALNSPR-----LVEDPL
NP_199131     (1)  -----MALEALNSPR-----LVEDPL
AF250336      (1)  -----MALEALTSPLASPIIP--P----LFEDSS
AY034998      (1)  -----MALEALTSPLASPIIP--P----LFEDSS
AY063006      (1)  -----MALEALTSPLASPIIP--P----LFEDSS
CAA64820      (1)  -----MALEALTSPLASPIIP--P----LFEDSS
NM_102538     (1)  -----MALEALTSPLASPIIP--P----LFEDSS
X95573        (1)  -----MALEALTSPLASPIIP--P----LFEDSS
CAA67228      (1)  -----MALEALTSPLASPIIP--P----LFEDSS
CAA67229      (1)  -----MALEALTSPLASPIIP--P----LFEDSS
X98670        (1)  -----MALEALTSPLASPIIP--P----LFEDSS
X98671        (1)  -----MALEALTSPLASPIIP--P----LFEDSS
BAC43454      (1)  -----MALETLSPLRSSPMP--T----LFQDSA
NP_196054     (1)  -----MALETLSPLRSSPMP--T----LFQDSA
BAA05079      (1)  -----MALEALNSPTTTTTPPS--F----QFENNG
BAA05076      (1)  -----MALEALKSPTAATPTLP-----PRYEDQV
BAA05077      (1)  -----MALEALKSPTAATPSLP-----PRYEDHV
BAA05078      (1)  -----MTLETLKSSSTPKTSKP-----TIPLPPK
NP_190562     (1)  -----MALDTLNSPTSTTTTT--APPPFLRCLDET
NP_201546     (1)  -----MALETLNSPTATTAR----PLLRREEM
NP_188592     (1)  -----MALEAMNTPTSSFTRIETKEDLMNDVAFIE
BAA19926      (276) QSENSDSEYFLGEYKKVESDASVDEFHRNGNYQWNTSNTSLGCWFDES GP
BAA20137      (288) QSESSDSAYFLEENAIVESDVSV DGFINGNSKWITSKMSNAAWCDESRT
BAA21919      (1)  -----MDLLQDRESETESLPYPTQCKRYKRIINSRISDTHYNQ
BAC43008      (63)  LEDG-----ESESERNVINLTRKRSKRTRKLD SFVTKVKVTS
NP_182037     (78)  LDDELEFDFAEDDDVESETESR--INPTRRRSKRTRKLG SFDFDFEKL T
AAL76091      (1)  -----
BAB67885      (1)  -----
CAA67233      (1)  -----MGQDEVGSDQTQIIKGKRTKRQRSSSTFVVTAATTVTSTSSSAGG
CAA67236      (1)  -----MGQDEVGSDQTQIIKGKRTKRQRSSSTFVVTAATTVTSTSSSAGG
BAA19111      (1)  -----MEALEEVVGVPNKDFHIIK GKRTKRLRLQSPIPFTVAATHDSSNA
BAA21920      (1)  -----MEVQM QEDHDDHMMNMVIKRRRTKRPRPSSP
CAA43111      (1)  -----MEFSEDSI
BAA19110      (1)  -----
BAA21923      (1)  -----
BAB67879      (1)  -----
BAA21921      (1)  -----
BAA21922      (1)  -----
CAA67231      (1)  -----
CAA67232      (1)  -----
CAA67234      (1)  -----
CAA67235      (1)  -----
CAB90936      (1)  -----
CAB90935      (1)  -----
CAA67230      (1)  -----
CAB67667      (1)  -----
BAA19114      (50)  DDDQASDWEDKFGGSVKEGNKRM YQIERTNPNRQKSNRVCENC GKEFSSWK

```

BAA21924 (1) -----
 BAA21925 (1) -----
 BAA21926 (1) -----
 CAA18741 (50) SSSS-----SSPKSASKPKYTKKPDNPAPKITRPCTECGRKF
 CAB80245 (50) SSSS-----SSPKSASKPKYTKKPDNPAPKITRPCTECGRKF
 Consensus (301) AL L

351 400

AAK01713 (35) ATSGEEGGHLPQG-----WAKRKRS--RRQR-----
 AF332876 (35) ATSGEEGGHLPQG-----WAKRKRS--RRQR-----
 AAM67193 (17) RFNGVEQWTKCKK-----RSKRSRSDLHHNHR-----
 NP_199131 (17) RFNGVEQWTKCKK-----RSKRSRSDLHHNHR-----
 AF250336 (24) VFHGVEHWTGK-----RSKRSRSDFHQ-----
 AY034998 (24) VFHGVEHWTGK-----RSKRSRSDFHQ-----
 AY063006 (24) VFHGVEHWTGK-----RSKRSRSDFHQ-----
 CAA64820 (24) VFHGVEHWTGK-----RSKRSRSDFHQ-----
 NM_102538 (24) VFHGVEHWTGK-----RSKRSRSDFHQ-----
 X95573 (24) VFHGVEHWTGK-----RSKRSRSDFHQ-----
 CAA67228 (24) VFHGVEHWTGK-----RSKRSRSDFHQ-----
 CAA67229 (24) VFHGVEHWTGK-----RSKRSRSDFHQ-----
 X98670 (24) VFHGVEHWTGK-----RSKRSRSDFHQ-----
 X98671 (24) VFHGVEHWTGK-----RSKRSRSDFHQ-----
 BAC43454 (24) LG---FHGSKGK-----RSKRSRSEFDRQS-----
 NP_196054 (24) LG---FHGSKGK-----RSKRSRSEFDRQS-----
 BAA05079 (24) LKY-LESWTGK-----RSKRQRS---MERQ-----
 BAA05076 (25) DMSNLDWVKGK-----RSKRPRITPPSE-----
 BAA05077 (25) DMNNLDWVKGK-----RSKRPRITPPSE-----
 BAA05078 (25) PINDAIDHKKR-----RSKRPRITPPSE-----
 NP_190562 (29) EPENLESWTGK-----RTKRHRIDQPNPP-----
 NP_201546 (26) EPENLEQWAKRK-----RTKRQRFDHGHQNETN-----
 NP_188592 (31) PWLKRKRKRQRS-----HSPSSSSSSPPSRPKS-----
 BAA19926 (326) AEKKELTRMKNYV-----TDSRKDLKDYKYDSYG-----
 BAA20137 (338) HKEKGLNRNKRNT-----IDSSKDITECEDDDYW-----
 BAA21919 (39) FLSLERRRQQQQQQ-----YGKITEFPFVESEP-----
 BAC43008 (102) QLGYPESDQEP-----HSSASDTTTEEDLAFCL-----
 NP_182037 (126) TSQPSELVAPEH-----HSSASDTTTEEDLAFCL-----
 AAL76091 (1) -----MTITREEAESK-----
 BAB67885 (1) -----
 CAA67233 (46) SGERAVSDEYNS-----AVSSPVTTDCTQEEE-----
 CAA67236 (46) SGERAVSDEYNS-----AVSSPVTTDCTQEEE-----
 BAA19111 (47) CDYHKNDTDNGEDVPHNNNNNNNDNNASPSSAPSAEMFIN-----
 BAA21920 (31) LALTIATSSCSTVEGTHAGELDGHVANSSSSPSNSG-----
 CAA43111 (9) DHTLVFKGKRKRPRQLSPDIYSSSTTSTTQISSSS-----
 BAA19110 (1) -----MVALSTKREREE-----
 BAA21923 (1) -----MVVLPLKRERE-----
 BAB67879 (1) -----MANCPSQNPATPAARRAPIAMSKRSRS-----
 BAA21921 (1) -----MSAMKRSR--E-----
 BAA21922 (1) -----MSSIKRSRSEE-----
 CAA67231 (1) -----MVAISE-----
 CAA67232 (1) -----MVAISE-----
 CAA67234 (1) -----MVARSEE-----
 CAA67235 (1) -----MVARSEE-----
 CAB90936 (1) -----MVARSEE-----
 CAB90935 (1) -----MVARSE-----
 CAA67230 (1) -----RSDFEES-----
 CAB67667 (1) -----MKRDRSDYEE-----
 BAA19114 (100) SFLEHGKCSSEDAEESLVSSPGSEGEDYIYDGRKEKGYGWSKRKRSRLRTK
 BAA21924 (1) -----MTIKRSREE-----
 BAA21925 (1) -----MTIKRSREE-----
 BAA21926 (1) -----MALEIRSRSEAK-----
 CAA18741 (87) WSWKALFGHMR-C-----HPERQWRGINPPPNYRV-----

CAB80245 (87) WSWKALFGHMR-C-----HPERQWRGINPPPNYRV-----
 Consensus (351) S---
 SEQ ID NO 8

401 450

AAK01713 (59) -----SEENLALCLMLLARGGHHR-----
 AF332876 (59) -----SEENLALCLMLLARGGHHR-----
 AAM67193 (44) -----LTEEEYLAFCLMLLARDGGD-----
 NP_199131 (44) -----LTEEEYLAFCLMLLARDGGD-----
 AF250336 (49) -----LTEEEYLAFCLMLLARDNRQ-----
 AY034998 (49) -----LTEEEYLAFCLMLLARDNRQ-----
 AY063006 (49) -----LTEEEYLAFCLMLLARDNRQ-----
 CAA64820 (49) -----LTEEEYLAFCLMLLARDNRQ-----
 NM_102538 (49) -----LTEEEYLAFCLMLLARDNRQ-----
 X95573 (49) -----LTEEEYLAFCLMLLARDNRQ-----
 CAA67228 (49) -----LTEEEYLAFWLMLLARDNRQ-----
 CAA67229 (49) -----LTEEEYLAFWLMLLARDNRQ-----
 X98670 (49) -----LTEEEYLAFWLMLLARDNRQ-----
 X98671 (49) -----LTEEEYLAFWLMLLARDNRQ-----
 BAC43454 (46) -----LTEDEYLALCLMLLARDGDRNRDLDP-----
 NP_196054 (46) -----LTEDEYLALCLMLLARDGDRNRDLDP-----
 BAA05079 (46) -----CTEEYIALCLIMLARS DGSVNNSRSL-----
 BAA05076 (50) -----EEYLALCLIMLARS GNGTTPSSIP-----
 BAA05077 (50) -----EEYLALCLIMLARS GNGTTPGSTD-----
 BAA05078 (50) -----KEFLALCLIMLARS G-GKNPTTTP-----
 NP_190562 (54) -----PSEEEYLALCLMLLARGSSDH-----
 NP_201546 (55) -----KNLPSSEEEYLALCLMLLARGSAVQ-----
 NP_188592 (61) ---Q-----NQDLTEEEYLALCLMLLAKDQPSQTRFHQQ-----
 BAA19926 (356) -----MVSHLDKRESRKRIKDSSYPNLSNETFKN-----
 BAA20137 (368) -----LSSYEDKCEPRKRERDSSYHPELGNESYKK-----
 BAA21919 (67) -----VSSISDTSDEEDVANCLMMLSRDKWMTQENEVI-----
 BAC43008 (132) -----MMLSRDKWKKNKS NKEVVEEIEETEE-----
 NP_182037 (156) -----IMLSRDKWKQOKKKQRVEEDET DHD-----
 AAL76091 (12) -----EMESLRVHASALLSLSSPAASAS-----
 BAB67885 (1) -----MESLRVHASALLSLSSPAASAS-----
 CAA67233 (74) -----DMAICLIMLARGTVLPS PDLKNSRKIH-----
 CAA67236 (74) -----DMAICLIMLARGTVLPS PDLKNSRKIH-----
 BAA19111 (87) -----TFTEDDEETARYLILMSK CDDHLHPRNKS RDGLPPDKFEL-----
 BAA21920 (67) -----IDILIRNREEDMANCLILLA QGHNNQKPSP-S-----
 CAA43111 (45) -----SREDEEDMANCLILLA QSGQSHKQKFSS-----
 BAA19110 (13) -----DNFYSITTMANYLMLLSRQANEHFDK-----
 BAA21923 (12) -----AEFKSITTMANYLMLLFS HQEN-HFNT-----
 BAB67879 (27) -----MWDMQEFVGSVD TARVLMLLAQSQHGLLGCGG-----
 BAA21921 (10) -----DRQVEAAAMANCLMMLLSKLNDKSTST-----
 BAA21922 (12) -----YGOVEAEAMANCLMMLLSKLNDHNTS-----
 CAA67231 (7) -----IKSTVDVTAANCLMMLLSRVGQENV D-----
 CAA67232 (7) -----IKSTVDVTAANCLMMLLSRVGQENV D-----
 CAA67234 (8) -----IVIVEEDTTAKCLMMLLSRVGECGGG-----
 CAA67235 (8) -----IVIVEEDTTAKCLMMLLSRVGECGGG-----
 CAB90936 (8) -----IVIVEEDTTAKCLMMLLSRVGECGGG-----
 CAB90935 (7) -----EVEIVEDTAKCLMMLLSRVGECGG-----
 CAA67230 (8) -----LKNIDI AKCLMILAQTS MVKQIGLNQ-----
 CAB67667 (12) -----MKHIDI VESLMMLSRSFVVKQIDVKQ-----
 BAA19114 (150) VGGLSTSTYQSSEEDLLAKCLIDLANARVDTSLVEPEESCASASREEE
 BAA21924 (11) -----LAKVETQATANCVNILEONTWLR-----
 BAA21925 (11) -----LAKVETLAMANCVNILEKNTSLAR-----
 BAA21926 (14) -----SEVSVERLAMENCANILQORNQLLGES-----
 CAA18741 (116) -----PTAASSKQLNQILPNWVS FMSEEDHEVASC-----
 CAB80245 (116) -----PTAASSKQLNQILPNWVS FMSEEDHEVASC-----
 Consensus (401) TEEELACLMLLAR

SEQ ID NO 9

| | 451 | 500 |
|-----------|-------|------------------------------------------------------|
| AAK01713 | (79) | -----VQA--PPPLSASAPPPAG-----AEFKCSVCGKSFSSYQAL |
| AF332876 | (79) | -----VQA--PPPLSASAPPPAG-----AEFKCSVCGKSFSSYQAL |
| AAM67193 | (64) | -----LDSVTVEEKP-----SYKCGVCYKTFSSYQAL |
| NP_199131 | (64) | -----LDSVTVAEKP-----SYKCGVCYKTFSSYQAL |
| AF250336 | (69) | -----PPPPPAVEKL-----SYKCSVCDKTFSSYQAL |
| AY034998 | (69) | -----PPPPPAVEKL-----SYKCSVCDKTFSSYQAL |
| AY063006 | (69) | -----PPPPPAVEKL-----SYKCSVCDKTFSSYQAL |
| CAA64820 | (69) | -----PPPPPAVEKL-----SYKCSVCDKTFSSYQAL |
| NM_102538 | (69) | -----PPPPPAVEKL-----SYKCSVCDKTFSSYQAL |
| X95573 | (69) | -----PPPPPAVEKL-----SYKCSVCDKTFSSYQAL |
| CAA67228 | (69) | -----PPPPPAVEKL-----SYKCSVCDKTFSSYQAL |
| CAA67229 | (69) | -----PPPPPAVEKL-----SYKCSVCDKTFSSYQAL |
| X98670 | (69) | -----PPPPPAVEKL-----SYKCSVCDKTFSSYQAL |
| X98671 | (69) | -----PPPPPAVEKL-----SYKCSVCDKTFSSYQAL |
| BAC43454 | (73) | -----SSSSS--PPLLPLPTP-----IYKCSVCDKAFSSYQAL |
| NP_196054 | (73) | -----SSSSS--PPLLPLPTP-----IYKCSVCDKAFSSYQAL |
| BAA05079 | (73) | -----PPPL--PPSVPTSQINATLLLEQKNLYKCSVCGKGFSSYQAL |
| BAA05076 | (74) | ---GSTDTTIS--KEPEKKNRDVAPVYQETEQS YKCSVCDKSFSSYQAL |
| BAA05077 | (74) | ---TTITTTIS--KEPEKKNRELTPVHQETEQS YKCSVCDKSFSSYQAL |
| BAA05078 | (73) | ---TTITNEPL--QVQEPINKPLQVQEPINEQS YKCNVCNKS FHSYQAL |
| NP_190562 | (75) | -----H--SPPSDHHSLSP LSDHOK--DYKCSVCGKSFSSYQAL |
| NP_201546 | (79) | -----S--PPLPLPSRASPSDHR--DYKCTVCGKSFSSYQAL |
| NP_188592 | (92) | -----SQ----SLTPPPESKN-----LPYKCNVCEKAFPSYQAL |
| BAA19926 | (386) | -----VKPSF--KSPEGSKHTQ-----KKKKYECLNCKKTFSSYQAL |
| BAA20137 | (398) | -----MKLSH--KGSEGCKKIH-----NKKKYECLNCKKIFGSSYQAL |
| BAA21919 | (100) | ---DNSASYDEDVKTEDSVVVKVTTTR--RGRGKYICETCNKVFRSSYQAL |
| BAC43008 | (158) | -----S--EGYNKINQAT-----TKGRYK CETCGKVFKSYQAL |
| NP_182037 | (182) | -----S--EDYKSS--K-----SRGRFKCETCGKVFKSYQAL |
| AAL76091 | (35) | -----QPTSSSST-----TEGVFECKTCSKRFPSEQAL |
| BAB67885 | (23) | -----QPTSSSST-----TEGVFECKTCSKRFPSEQAL |
| CAA67233 | (101) | -----QKISSENS-----SFYVFECKTCNRTFSSSEQAL |
| CAA67236 | (101) | -----QKISSENS-----SFYVFECKTCNRTFSSSEQAL |
| BAA19111 | (127) | FNDDLKLYQSKFNSKRYIETSTNLGNGTKAGIFVFECKTCNRTFPSEQAL |
| BAA21920 | (99) | -----HSPLDVYQCKTCNRCFPSEQAL |
| CAA43111 | (73) | -----RKFTETATSTGKAGFYVFECKTCNRTFPSEQAL |
| BAA19110 | (39) | -----KMNNSS-----TSRVFECKTCNRCFPSEQAL |
| BAA21923 | (37) | -----MMDNS-----PSRVFECKTCNRCFPSEQAL |
| BAB67879 | (60) | -----FAAGAQPVVVRGGAHDRVFECKTCNRCFPSEQAL |
| BAA21921 | (36) | -----TTTNQD-----HHNDFECKTCNKRFPSEQAL |
| BAA21922 | (37) | -----KNQD-----HHNEFECKTCNKRFPSEQAL |
| CAA67231 | (32) | -----GGD-----QKRVECTCKTCLKQFHSSEQAL |
| CAA67232 | (32) | -----GGD-----QKRVECTCKTCLKQFHSSEQAL |
| CAA67234 | (33) | -----CGG-----DERVFRCKTCLKEFSSSEQAL |
| CAA67235 | (33) | -----CGG-----DERVFRCKTCLKEFSSSEQAL |
| CAB90936 | (33) | -----CGG-----DERVFRCKTCLKEFSSSEQAL |
| CAB90935 | (31) | -----GG-----EKRVFRCKTCLKEFSSSEQAL |
| CAA67230 | (34) | -----HTESHT-----SNQFECKTCNKRFPSEQAL |
| CAB67667 | (38) | -----STGSKTN-----HNNHFECKTCNKRFPSEQAL |
| BAA19114 | (200) | RAARNSMAYGFTPLVSTRVPFDNKA KGAS-SKGLFECKACKKVFN SHQAL |
| BAA21924 | (35) | -----KIFECKTCKKQFDSEQAL |
| BAA21925 | (35) | -----KIFECKTCKKQFDSEQAL |
| BAA21926 | (41) | -----S-----SKIFECKTCKKQFDSEQAL |
| CAA18741 | (146) | -----LLMLS--NGTPSS--S-----SIERFECGCKKVFGSHQAL |
| CAB80245 | (146) | -----LLMLS--NGTPSS--S-----SIERFECGCKKVFGSHQAL |
| Consensus | (451) | YKC C K FSSYQAL |

SEQ ID NO 5-firstpart

| | 501 | 550 |
|-----------|-----------------------------------------------------------|-------------|
| AAK01713 | (113) GGHKTSHRVKLPT----- | PPAAP |
| AF332876 | (113) GGHKTSHRVKLPT----- | PPAAP |
| AAM67193 | (91) GGHKASHRS----- | LYGG |
| NP_199131 | (91) GGHKASHRS----- | LYGG |
| AF250336 | (96) GGHKASHRK-NLSQ----- | TLSGG |
| AY034998 | (96) GGHKASHRK-NLSQ----- | TLSGG |
| AY063006 | (96) GGHKASHRK-NLSQ----- | TLSGG |
| CAA64820 | (96) GGHKASHRK-NLSQ----- | TLSGG |
| NM_102538 | (96) GGHKASHRK-NLSQ----- | TLSGG |
| X95573 | (96) GGHKASHRK-NLSQ----- | TLSGG |
| CAA67228 | (96) GGHKASHRK-NLSQ----- | TLSGG |
| CAA67229 | (96) GGHKASHRK-NLSQ----- | TLSGG |
| X98670 | (96) GGHKASHRK-NLSQ----- | TLSGG |
| X98671 | (96) GGHKASHRK-NLSQ----- | TLSGG |
| BAC43454 | (105) GGHKASHRK-SFSL----- | TQSAG |
| NP_196054 | (105) GGHKASHRK-SFSL----- | TQSAG |
| BAA05079 | (114) GGHKASHRK----- | LVSMS |
| BAA05076 | (119) GGHKASHRK-IITI----- | ATTAL |
| BAA05077 | (118) GGHKASHRK-IITI----- | ATTAL |
| BAA05078 | (117) GGHKASHRK-NKLS----- | TTTVS |
| NP_190562 | (110) GGHKTSHRKPVSV----- | DVNNNS |
| NP_201546 | (113) GGHKTSHRKPTNT----- | SITSG |
| NP_188592 | (122) GGHKASHRIKPPT----- | VIST |
| BAA19926 | (421) GGHRPCNKNTNAYFESTYETGENSRDADNGPNYINKGKHRETLSNKPAAH | |
| BAA20137 | (433) GGHRPCHKKANSYVESINGTGENSLDADHDG----- | KPFSA |
| BAA21919 | (145) GGHRASHKKIKVS----- | |
| BAC43008 | (189) GGHRASHKKNRVSNNKTEQR----- | |
| NP_182037 | (210) GGHRASHKKNKACMTKTEQ----- | |
| AAL76091 | (63) GGHRTSHTR----- | LQAKL |
| BAB67885 | (51) GGHRTSHTR----- | LQAKL |
| CAA67233 | (129) GGHRASHKKPRTSTEEKTRLPLTPQKSSASEEG----- | QNSHFKV |
| CAA67236 | (129) GGHRASHKKPRTSTEEKTRLPLTPQKSSASEEG----- | QNSHFKV |
| BAA19111 | (177) GGHRASHKKPKTLTTELVNKK----- | LYFDFS |
| BAA21920 | (121) GGHRASHKKPKLPTNLEEKNSKPIEHVENCSSKSNEDHVTTLSQLISNNNI | |
| CAA43111 | (106) GGHRTSHKKSKTIAAEKTSTLEDHQQQERVAQ----- | EEGEFIKIIPS |
| BAA19110 | (65) GGHRASHKKPRL----- | M |
| BAA21923 | (62) GGHRASHKKPRL----- | M |
| BAB67879 | (94) GGHRASHKKPRQQ-----QQ----- | HALGGG |
| BAA21921 | (62) GGHRASHKKPK----- | |
| BAA21922 | (61) GGHRASHKKTK----- | |
| CAA67231 | (55) GGHRASHKKPN----- | |
| CAA67232 | (55) GGHRASHKKPN----- | |
| CAA67234 | (56) GGHRASHKKLIN----- | S |
| CAA67235 | (56) GGHRASHKKLIN----- | S |
| CAB90936 | (56) GGHRASHKKLIN----- | S |
| CAB90935 | (53) GGHRASHKKLIN----- | S |
| CAA67230 | (59) GGHRASHKKPKL----- | T |
| CAB67667 | (65) GGHRASHKKPKL----- | I |
| BAA19114 | (249) GGHRASHKKVKG----- | CYAA |
| BAA21924 | (53) GGHRASHKKPK----- | F |
| BAA21925 | (53) GGHRASHKKPR----- | F |
| BAA21926 | (61) GGHRTSHKILRN----- | KL |
| CAA18741 | (178) GGHRASHKKNVKGCFAITNVTDGP----- | MT |
| CAB80245 | (178) GGHRASHKKNVKGCFAITNVTDGP----- | MT |
| Consensus | (501) GGHRASHKK | |

SEQ ID NO 5-cont

| | | 551 | 600 |
|-----------|-------|----------------------------------------|--------------------------|
| AAK01713 | (131) | VLAPAPVAALLPSAEDREPATSSTAASSDGMTNRVHRC | SICQKEFPTGQA |
| AF332876 | (131) | VLAPAPVAALLPSAEDREPATSSTAASSDGMTNRVHRC | SICQKEFPTGQA |
| AAM67193 | (104) | GDNDKSTPSTA-----V----- | KSHVCSVCGKSFATGQA |
| NP_199131 | (104) | GENDKSTPSTA-----V----- | KSHVCSVCGKSFATGQA |
| AF250336 | (114) | GDDHSTSSAITTTSA-----VTTGSG---- | KSHVCTICNKSFPSSGQA |
| AY034998 | (114) | GDDHSTSSAITTTSA-----VTTGSG---- | KSHVCTICNKSFPSSGQA |
| AY063006 | (114) | GDDHSTSSAITTTSA-----VTTGSG---- | KSHVCTICNKSFPSSGQA |
| CAA64820 | (114) | GDDHSTSSAITTTSA-----VTTGSG---- | KSHVCTICNKSFPSSGQA |
| NM_102538 | (114) | GDDHSTSSAITTTSA-----VTTGSG---- | KSHVCTICNKSFPSSGQA |
| X95573 | (114) | GDDHSTSSAITTTSA-----VTTGSG---- | KSHVCTICNKSFPSSGQA |
| CAA67228 | (114) | GDDHSTSSAITTTSA-----VTTGSG---- | KSHVCTICNKSFPSSGQA |
| CAA67229 | (114) | GDDHSTSSAITTTSA-----VTTGSG---- | KSHVCTICNKSFPSSGQA |
| X98670 | (114) | GDDHSTSSAITTTSA-----VTTGSG---- | KSHVCTICNKSFPSSGQA |
| X98671 | (114) | GDDHSTSSAITTTSA-----VTTGSG---- | KSHVCTICNKSFPSSGQA |
| BAC43454 | (123) | GDELSTSSAITTSG-----ISGGGGGGSV-- | KSHVCSICHKSFATGQA |
| NP_196054 | (123) | GDELSTSSAITTSG-----ISGGGGGGSV-- | KSHVCSICHKSFATGQA |
| BAA05079 | (128) | GDEQSTTSTTTNVT-----GTSSANVNGNGR | THECSICHKCFPTGQA |
| BAA05076 | (137) | LDDNNNNPTTSNSTNG---NVVNNISTLNP | SGRSHVCSICHKAFPSGQA |
| BAA05077 | (136) | LDDNNNNPTTSNSTSG---NVVNNISALNP | SGRSHVCSICHKAFPSGQA |
| BAA05078 | (135) | YDDTN--PSTSNSLN-----PSGRFHEC | SICHKCFPSGQA |
| NP_190562 | (128) | NGTVTNNGNISNG-----LVGQSGKTHNC | SICFKSFPSSGQA |
| NP_201546 | (131) | NQELSNNSHNSGSGVV--INVTVNTGNGV | SQSGKIHTCSICFKSFASGQA |
| NP_188592 | (139) | TADDSTAPTISIVAG-----EKHPAAASCK | IEHC SICHKVFPPTGQA |
| BAA19926 | (471) | SHDYSSNPEKKMKP-----KKFKGHAC | PFCPPMPKSGQA |
| BAA20137 | (469) | VKEPSYNPEKKIKP-----KKVKGHEC | PCYCDRVFKSGQA |
| BAA21919 | (158) | --INETKNNGNVES-----EVQKDKI | HECPVCYRVFSSGQA |
| BAC43008 | (209) | ----SETEYDNVVV-----VAKRI | HECPICLRVVFASGQA |
| NP_182037 | (229) | ----VETEVVLGVK-----EKK-V | HECPICFRVFTSGQA |
| AAL76091 | (77) | LSDPAAAAAAAAAER-----DRARV | HECAVCGVEFSMGQA |
| BAB67885 | (65) | LSDPAAAAAAAAAER-----DRARV | HECAVCGVEFSMGQA |
| CAA67233 | (169) | SGSALASQASNIIN-----KANKV | HECSICGSEFTSGQA |
| CAA67236 | (169) | SGSALASQASNIIN-----KANKV | HECSICGSEFTSGQA |
| BAA19111 | (205) | DEDDQSPSSTLCKTKNDVNRILPNSSNKYTSP | RIHECSYCGAEFTSGQA |
| BAA21920 | (171) | NNNNNSNNNNNNNIK-----NKNRV | HECSICGAEFTSGQA |
| CAA43111 | (150) | ISTQIINKGNMQSNFN-----SKSKI | HECATCGAEFTSGQA |
| BAA19110 | (78) | GELHNLQLFHE--L-----PKRK | THECSICGLEFAIGQA |
| BAA21923 | (75) | GELN-FQLPTS--P-----PKPK | THECSICGLEFPDGGQA |
| BAB67879 | (115) | AGADDAGLCLGRQP-----TPPRPQPAKPR | VHECPVCGLEFPDGGQA |
| BAA21921 | (73) | -LLIGAGEFLVQPS-----SK-KM | HECSICGMEFSLGQA |
| BAA21922 | (72) | -VLTGAGEFLAQQA-----KKNKM | HECSICGMEFSLGQA |
| CAA67231 | (66) | -NDALSSGLMKK-----VKTSS | HPCPICGVEFPMGQA |
| CAA67232 | (66) | -NDALSSGLMKK-----VKTSS | HPCPICGVEFPMGQA |
| CAA67234 | (69) | DNPSLLGSLSNK-----KTKTSH | PCPICGVKFPMGQA |
| CAA67235 | (69) | DNPSLLGSLSNK-----KTKTSH | PCPICGVKFPMGQA |
| CAB90936 | (69) | DNPSLLGSLSNK-----KTKTSH | PCPICGVKFPMGQA |
| CAB90935 | (66) | SDPSLLGSLSNKKT-----KTATSH | PCPICGVVEFPMGQA |
| CAA67230 | (72) | VEQKDVKHLND-Y-----KGNHF | HKCSICSQSFGTGA |
| CAB67667 | (78) | VDQEQVKHRN-----KENDMH | KCTICDOMFGTGA |
| BAA19114 | (265) | KQDQLDDILIDDQDVNITHDQEF | LQSSKSMRKS KIHECSICHKVFS |
| BAA21924 | (65) | ITAADFSIETSSY-----EC | SFCGEDFPTGQA |
| BAA21925 | (65) | ITAADFSICSPNYK-----PNNK | VHECSFCGEDFPTGQA |
| BAA21926 | (75) | LTSPLPGNDQLPVKT-----KK | HECSICGEQFLLGQA |
| CAA18741 | (203) | VSTSSGHDHQKIL-----TFSGH | HKCNICFVVFSSGQA |
| CAB80245 | (203) | VSTSSGHDHQKIL-----TFSGH | HKCNICFVVFSSGQA |
| Consensus | (551) | T | K H CSIC K FPSGQA |

SEQ ID NO 5-firstpart

| | | 601 | 650 |
|-----------|-------|-----------------------------------------|-------------------------|
| AAK01713 | (181) | LGGHKRCHYDGGVGA-----GAGASSTELLATVAAESEV | GSSGNGQS |
| AF332876 | (181) | LGGHKRCHYDGGVGA-----GAGASSTELLATVAAESEV | GSSGNGQS |
| AAM67193 | (133) | LGGHKRCHYDGGVS-----NSE---- | GVGSTSHVSSSS---- |
| NP_199131 | (133) | LGGHKRCHYDGGVS-----NSE---- | GVGSTSHVSSSS---- |
| AF250336 | (151) | LGGHKRCHYEGNNN-----INTSSVSNSEGA | STSHVSSS---- |
| AY034998 | (151) | LGGHKRCHYEGNNN-----INTSSVSNSEGA | STSHVSSS---- |
| AY063006 | (151) | LGGHKRCHYEGNNN-----INTSSVSNSEGA | STSHVSSS---- |
| CAA64820 | (151) | LGGHKRCHYEGNNN-----INTSSVSNSEGA | STSHVSSS---- |
| NM_102538 | (151) | LGGHKRCHYEGNNN-----INTSSVSNSEGA | STSHVSSS---- |
| X95573 | (151) | LGGHKRCHYEGNNN-----INTSSVSNSEGA | STSHVSSS---- |
| CAA67228 | (151) | LGGHKRCHYEGNNN-----INTSSVSNSEGA | STSHVSSS---- |
| CAA67229 | (151) | LGGHKRCHYEGNNN-----INTSSVSNSEGA | STSHVSSS---- |
| X98670 | (151) | LGGHKRCHYEGNNN-----INTSSVSNSEGA | STSHVSSS---- |
| X98671 | (151) | LGGHKRCHYEGNNN-----INTSSVSNSEGA | STSHVSSS---- |
| BAC43454 | (163) | LGGHKRCHYEGKNGG-----GVSSSVNS | SEDVGSTSHVSSG---- |
| NP_196054 | (163) | LGGHKRCHYEGKNGG-----GVSSSVNS | SEDVGSTSHVSSG---- |
| BAA05079 | (170) | LGGHKRCHYDGGNG-----NNGSVSV | GVGTSSEGVG |
| BAA05076 | (183) | LGGHKRCHYEGKLCG-----NNNNNR | DCGCHSGSVVTTSDGGASTH |
| BAA05077 | (182) | LGGHKRCHYEGKLCG-----NNNN-NR | DGCHSGSVVTTSDGGASTH |
| BAA05078 | (168) | LGGHKRCHYEGNLGG-----GVSR----- | GD--TVISSEGGGSAV |
| NP_190562 | (164) | LGGHKRCHYDGGN-----GNSNGD----- | NSH---- |
| NP_201546 | (179) | LGGHKRCHYDGGN-----GNGNGS----- | SSN---- |
| NP_188592 | (180) | LGGHKRCHYEGNLGGG-----GGGSKS | ISHSGSVSTVSEERS--- |
| BAA19926 | (505) | LGGHKRSHYIVSSE-----NHYQAS | AVQVGK |
| BAA20137 | (503) | LGGHKRSHYIGGSFR-----NLNQSS | AAKKE |
| BAA21919 | (192) | LGGHKRSHGIGVAATN-----VSLSTK | IVSSRI |
| BAC43008 | (239) | LGGHKRSHGVGN-----LSVN----- | QQRVRHRESV |
| NP_182037 | (258) | LGGHKRSHGSGNIGAGRG-----LSVS----- | QIQIEEEVSV |
| AAL76091 | (111) | LGGHMRHRGETGTTTV-----VLADADD | SG-----GATVPQP- |
| BAB67885 | (99) | LGGHMRHRGETGTTTV-----VLADADD | SG-----GATVPQP- |
| CAA67233 | (203) | LGGHMRHRHTAVTTISP-----VAATAE | VS RNSTEEIEIINIGRSMEQQR |
| CAA67236 | (203) | LGGHMRHRHTAVTTISP-----VAATAE | VS RNSTEEIEIINIGRSMEQQR |
| BAA19111 | (255) | LGGHMRHRHGGVNVN-----SSLHLS | NYVSPATSIDQEFANNTMKKVP |
| BAA21920 | (206) | LGGHMRHRHPLPS-----IAIAST | SHELESSHEIKNTR |
| CAA43111 | (187) | LGGHMRHRPPTITANITNTKVTLS | TIDDTSNYTSSESSH |
| BAA19110 | (110) | LGGHMRHRRAVIN-----DKN-LQ | APDDQ----HAPVVKKANGRR |
| BAA21923 | (106) | LGGHMRHRRAVMN-----ENN-LQ | VT-----PVVKKNSRR |
| BAB67879 | (156) | LGGHMRHRRAEAAEAAAT-----TTTTTT | TKNGDVGKAAAVKACDG |
| BAA21921 | (105) | LGGHMRHRRAAIDE-----KSK-- | AATK-AM----MIPVLKKSNS |
| BAA21922 | (105) | LGGHMRHRRDENNK-----TLK-- | VARKTTT----MIPVLKKSNS |
| CAA67231 | (97) | LGGHMRHRRNESGAAG-----ALVTR | ALLPEP-----TVTTLKKSSSGK |
| CAA67232 | (97) | LGGHMRHRRNESGAAG-----ALVTR | ALLPEP-----TVTTLKKSSSGK |
| CAA67234 | (101) | LGGHMRHRNEKVS--G-----SLVTR | SFLPETT----TVTALKKFSSGK |
| CAA67235 | (101) | LGGHMRHRNEKVS--G-----SLVTR | SFLPETT----TVTALKKFSSGK |
| CAB90936 | (101) | LGGHMRHRNEKVS--G-----SLVTR | SFLPETT----TVTALKKFSSGK |
| CAB90935 | (100) | LGGHMRHRSEKASP-G-----TLVTR | SFLPETT----TVTTLKKSSSGK |
| CAA67230 | (105) | LGGHMRHRSSMTV-----EPS-FIS | PMIP----SMPVLKRCGSSK |
| CAB67667 | (108) | LGGHMRHRRTSMIT-----EQS-IV | PSVY----SRPVFNRCSSSK |
| BAA19114 | (315) | LGGHKRCHWITSNS-----PDSSK | FHFNGHVEQINLRSNMH |
| BAA21924 | (92) | LGGHMRKHPDKL-----KKKKQK | NCDDKLGSGKS----- |
| BAA21925 | (99) | LGGHMRKHRDKLGH-----EL-KK | QKQKNCDEIVEIEKKNNS---GT |
| BAA21926 | (106) | LGGHMRKHRDELNQ-----LQQQKK | KIKMDDEKSDVSEEVVQEKGN |
| CAA18741 | (237) | LGGHMRCHWEKE-----EPMISG--- | |
| CAB80245 | (237) | LGGHMRCHWEKE-----EPMISG--- | |
| Consensus | (601) | LGGHKRRHYEG | T V |

SEQ ID NO 5-CONT

| | | 651 | 700 |
|-----------|-------|-----------------------------------------|--------------------|
| AAK01713 | (223) | ATRAFDLNLPAVPEFVWRPCSKGK--KMWD---- | EEEEVQSPLAFKKPRL |
| AF332876 | (223) | ATRAFDLNLPAVPEFVWRPCSKGK--KMWD---- | EEEEVQSPLAFKKPRL |
| AAM67193 | (162) | -HRGFDLNLIPVQGFSR----- | DDEVMSPMATKKPRL |
| NP_199131 | (162) | -HRGFDLNLIPVQGFSR----- | DDEVMSPMATKKPRL |
| AF250336 | (186) | -HRGFDLNLIPPTPEFSMVNG----- | DDEVMSMPAKKPRF |
| AY034998 | (186) | -HRGFDLNLIPPTPEFSMVNG----- | DDEVMSMPAKKPRF |
| AY063006 | (186) | -HRGFDLNLIPPTPEFSMVNG----- | DDEVMSMPAKKPRF |
| CAA64820 | (186) | -HRGFDLNLIPPTPEFSMVNG----- | DDEVMSMPAKKPRF |
| NM_102538 | (186) | -HRGFDLNLIPPTPEFSMVNG----- | DDEVMSMPAKKPRF |
| X95573 | (186) | -HRGFDLNLIPPTPEFSMVNG----- | DDEVMSMPAKKPRF |
| CAA67228 | (186) | -HRGFDLNLIPPTPEFSMVNG----- | DDEVMSMPAKKPRF |
| CAA67229 | (186) | -HRGFDLNLIPPTPEFSMVNG----- | DDEVMSMPAKKPRF |
| X98670 | (186) | -HRGFDLNLIPPTPEFSMVNG----- | DDEVMSMPAKKPRF |
| X98671 | (186) | -HRGFDLNLIPPTPEFSMVNG----- | DDEVMSMPAKKPRF |
| BAC43454 | (199) | -HRGFDLNLIPPTPEFSMVNG----- | DEEVMSMPAKKPRF |
| NP_196054 | (199) | -HRGFDLNLIPPTPEFSMVNG----- | DEEVMSMPAKKPRF |
| BAA05079 | (206) | -HRGFDLNLIPALPEFWPGFGSG----- | EDEVSPHAKKPRF |
| BAA05076 | (225) | TLRDFDLNMLEPSPQLGLSIDC--GLKSQQVPIEQEVES | SPMLPKKPRF |
| BAA05077 | (223) | TLRDFDLNMLEPSPQLGLSIDC--DLKS-QIPIEQEVES | SPMLPKKPRF |
| BAA05078 | (201) | IRRDFDLNMLEPSPQLGLSIDC--ERKS-QLSGEQEVES | SPMPTKKPRF |
| NP_190562 | (186) | ---KFDLNLPAQVSDDETIGKS----- | Q---LSGEETKSVL- |
| NP_201546 | (202) | ---SVELVAGSDVSDVDNERWS----- | EESAIGGHRGFDLNL |
| NP_188592 | (219) | HRGFDLNLPAPELSLHHN----- | PIVDEEILSPITGKKPLL |
| BAA19926 | (530) | IVDLIDLNLPAPELVDVNGE----- | PAFVSW----- |
| BAA20137 | (529) | ADDLIDLNLPAPELDEDNEH----- | AHFVSW----- |
| BAA21919 | (220) | SGTMIDLNLPAPELDEISO----- | IEVSAVSDDEFVNP |
| BAC43008 | (266) | KQRMIDLNLPAPELDEEVS----- | VVFQ----- |
| NP_182037 | (290) | KQRMIDLNLPAPELDEETS----- | LVFDEW----- |
| AAL76091 | (144) | -----PEPMPDLNYP----- | LEDAGDGSEPELL |
| BAB67885 | (132) | -----PEPMPDLNYP----- | LEDAGDGSEPELL |
| CAA67233 | (248) | KYLPIDLNLPAPELDELRISK----- | FQGI VFSATPALI |
| CAA67236 | (248) | KYLPIDLNLPAPELDELRISK----- | FQGI VFSATPALI |
| BAA19111 | (299) | DGLSIDLNLPAPELDELRISK----- | LKQOQDEKTQROQLV |
| BAA21920 | (240) | NFLSIDLNLPAPELDELRISK----- | KEQVIVFSASPLVD |
| CAA43111 | (237) | IILSIDLNLPAPELDELRISK----- | GNKQCLVFSAAALV |
| BAA19110 | (145) | -ILSIDLNLPAPELDELRISK----- | LRKSNTPMVDCL |
| BAA21923 | (136) | -VLCLIDLNLPAPELDELRISK----- | LGK--AARIVDCL |
| BAB67879 | (195) | GGVCLIDLNLPAPELDELRISK----- | LGAGGQGVHAKALML |
| BAA21921 | (140) | RIFCLIDLNLPAPELDELRISK----- | WPTAPITSSPVLRI |
| BAA21922 | (141) | RIFCLIDLNLPAPELDELRISK----- | WPTAPITSSPVLRI |
| CAA67231 | (137) | RVACIDLNLPAPELDELRISK----- | ELGRTIVY----- |
| CAA67232 | (137) | RVACIDLNLPAPELDELRISK----- | ELGRTIVY----- |
| CAA67234 | (140) | RVACIDLNLPAPELDELRISK----- | ELGRTISWS---- |
| CAA67235 | (140) | RVACIDLNLPAPELDELRISK----- | ELGRTISWS---- |
| CAB90936 | (140) | RVACIDLNLPAPELDELRISK----- | ELGRTISWS---- |
| CAB90935 | (140) | RVACIDLNLPAPELDELRISK----- | ELGRTISWS---- |
| CAA67230 | (141) | RILSIDLNLPAPELDELRISK----- | FGKTFVPKIDMKF |
| CAB67667 | (144) | --EILIDLNLPAPELDELRISK----- | FGKNLVPQIDLKF |
| BAA19114 | (351) | KSDALIDLNLPAPELDELRISK----- | DIHLQYPWSCAPKNDN |
| BAA21924 | (121) | --LFLIDLNLPAPELDELRISK----- | IPLKVFD----- |
| BAA21925 | (137) | GKLFFIDLNLPAPELDELRISK----- | IPLKVFI----- |
| BAA21926 | (148) | AGLFFIDLNLPAPELDELRISK----- | TSVHS----- |
| CAA18741 | (256) | ---ALDLNVPPTIQDLSTSD----- | TSGCCIDLRLGL- |
| CAB80245 | (256) | ---ALDLNVPPTIQDLSTSD----- | TSGCCIDLRLGL- |
| Consensus | (651) | LDLNLPP I ED | EV S M |

SEQ ID NO 7

| | 701 | 750 |
|-----------|----------------------------------------------------------|-----|
| AAK01713 | (267) LTA----- | |
| AF332876 | (267) LTA----- | |
| AAM67193 | (193) K----- | |
| NP_199131 | (193) K----- | |
| AF250336 | (220) DFPVKLQL----- | |
| AY034998 | (220) DFPVKLQL----- | |
| AY063006 | (220) DFPVKLQL----- | |
| CAA64820 | (220) DFPVKLQL----- | |
| NM_102538 | (220) DFPVKLQL----- | |
| X95573 | (220) DFPVKLQL----- | |
| CAA67228 | (220) DFPVKLQL----- | |
| CAA67229 | (220) DFPVKLQL----- | |
| X98670 | (220) DFPVKLQL----- | |
| X98671 | (220) DFPVKLQL----- | |
| BAC43454 | (233) DFPGKP----- | |
| NP_196054 | (233) DFPEKP----- | |
| BAA05079 | (242) SLPPKLELFKGL----- | |
| BAA05076 | (273) LFSMD----- | |
| BAA05077 | (270) LFSMD----- | |
| BAA05078 | (247) AFRIDGN----- | |
| NP_190562 | (216) ----- | |
| NP_201546 | (236) PADQVSVTTS----- | |
| NP_188592 | (257) LTDHDQVIKKEDLSLKI----- | |
| BAA19926 | (555) ----- | |
| BAA20137 | (555) ----- | |
| BAA21919 | (254) ----- | |
| BAC43008 | (289) ----- | |
| NP_182037 | (315) ----- | |
| AAL76091 | (168) NLLV----- | |
| BAB67885 | (156) NLLV----- | |
| CAA67233 | (281) DCHY----- | |
| CAA67236 | (281) DCHY----- | |
| BAA19111 | (338) DCHY----- | |
| BAA21920 | (280) CHY----- | |
| CAA43111 | (278) DCHY----- | |
| BAA19110 | (176) L----- | |
| BAA21923 | (166) L----- | |
| BAB67879 | (233) DCFL----- | |
| BAA21921 | (173) FI----- | |
| BAA21922 | (174) FI----- | |
| CAA67231 | (163) ----- | |
| CAA67232 | (163) ----- | |
| CAA67234 | (169) ----- | |
| CAA67235 | (169) ----- | |
| CAB90936 | (169) ----- | |
| CAB90935 | (165) ----- | |
| CAA67230 | (173) VL----- | |
| CAB67667 | (174) VN----- | |
| BAA19114 | (401) DNYYLEEIKIDSNANNGKYNINNGATQNVEDDEADSKLKLAKLSDLKDMN | |
| BAA21924 | (147) ----- | |
| BAA21925 | (164) ----- | |
| BAA21926 | (173) ----- | |
| CAA18741 | (285) ----- | |
| CAB80245 | (285) ----- | |
| Consensus | (701) | |

| | 751 | 774 |
|-----------|-------|---------------------------|
| AAK01713 | (270) | ----- |
| AF332876 | (270) | ----- |
| AAM67193 | (194) | ----- |
| NP_199131 | (194) | ----- |
| AF250336 | (228) | ----- |
| AY034998 | (228) | ----- |
| AY063006 | (228) | ----- |
| CAA64820 | (228) | ----- |
| NM_102538 | (228) | ----- |
| X95573 | (228) | ----- |
| CAA67228 | (228) | ----- |
| CAA67229 | (228) | ----- |
| X98670 | (228) | ----- |
| X98671 | (228) | ----- |
| BAC43454 | (239) | ----- |
| NP_196054 | (239) | ----- |
| BAA05079 | (254) | ----- |
| BAA05076 | (278) | ----- |
| BAA05077 | (275) | ----- |
| BAA05078 | (254) | ----- |
| NP_190562 | (216) | ----- |
| NP_201546 | (246) | ----- |
| NP_188592 | (274) | ----- |
| BAA19926 | (555) | ----- |
| BAA20137 | (555) | ----- |
| BAA21919 | (254) | ----- |
| BAC43008 | (289) | ----- |
| NP_182037 | (315) | ----- |
| AAL76091 | (172) | ----- |
| BAB67885 | (160) | ----- |
| CAA67233 | (285) | ----- |
| CAA67236 | (285) | ----- |
| BAA19111 | (342) | ----- |
| BAA21920 | (283) | ----- |
| CAA43111 | (282) | ----- |
| BAA19110 | (177) | ----- |
| BAA21923 | (167) | ----- |
| BAB67879 | (237) | ----- |
| BAA21921 | (175) | ----- |
| BAA21922 | (176) | ----- |
| CAA67231 | (163) | ----- |
| CAA67232 | (163) | ----- |
| CAA67234 | (169) | ----- |
| CAA67235 | (169) | ----- |
| CAB90936 | (169) | ----- |
| CAB90935 | (165) | ----- |
| CAA67230 | (175) | ----- |
| CAB67667 | (176) | ----- |
| BAA19114 | (451) | TNSDNPAAHWLQVGIGSTTEVGADS |
| BAA21924 | (147) | ----- |
| BAA21925 | (164) | ----- |
| BAA21926 | (173) | ----- |
| CAA18741 | (285) | ----- |
| CAB80245 | (285) | ----- |
| Consensus | (751) | |

Annex 5

AtSTZparalog2: SEQ ID NO: 28 under the control of the beta-expansine EXPB9 (PRO061_2)

A DNA fragment encoding a 2XC2H2 protein represented in the application as filed by SEQ ID NO: 29 was isolated from an *Arabidopsis thaliana* seedling cDNA library (Invitrogen, Paisley, UK) by PCR amplification and subsequent cloning in an entry clone vector according to the methods described in Example 1 of the present application. SEQ ID NO: 29 polypeptide was encoded by the longest open reading frame of SEQ ID NO: 28 (AtSTZparalog2).

The primers used for the PCR amplification were as follows:

Forward primer:

ggggacaagtttgtacaaaaaagcaggcttaaacaatggcacttgaaactcttact

Reverse primer:

ggggaccactttgtacaagaaagctgggtttcctaggtttatgttttaggg

The entry clone was subsequently used in an LR reaction with a destination vector used for rice transformation to generate the plant expression vector pEXPB9::AtSTZparalog2. The PCR amplified DNA fragmented was operably linked to the promoter PRO061_2 (pEXPB9: beta-expansine EXPB9), which is described in Table 10 in Table 10 on page 48 of the application as filed.

Phenotypic characterization of the transformed plants was carried out essentially as described in Example 3 of the present application. The results are shown in Table I below.

Table I: Results of phenotypic characterization of T1 rice plants transformed with pEXPB9::AtSTZparalog2.

| pEXPB9::AtSTZparalog2 | |
|------------------------|-----------------------------------------------------------------------|
| Parameter | % increase in the transgenic plants compared to the nullzygous plants |
| Number of filled seeds | 6.2 |
| Seed filing rate | 3.4 |

| | |
|---------------------------|-----|
| Harvest index | 3.9 |
| Number of panicles | 3.8 |

The above results show that overexpression of the nucleic acid represented by SEQ ID NO: 28 (encoding the 2XC2H2 zinc finger protein represented by SEQ ID NO 29) under the control of a promoter expressed in young expanding tissues (pEXP9) gives an increase in plant yield as shown by the increased in the number of filled seeds, the increase in the seed filling rate (proportion of filled seed with respect to the total number of seeds harvested per plant, increased in the harvest index and the increase in the number of panicles per plant.

Annex 6

SEQ ID NO: 7 XDLNXXP

Table 1. Percentage identity of the EAR motif as found in the in the amino acid of the sequence listing of the application as filed and SEQ ID NO 7.

| | Motif | % Identity to SEQ ID NO: 7 (XDLNXXP) | % Identity to SEQ ID NO: 7 XDLNXP |
|---------------------------------|---------|-----------------------------------------|--------------------------------------|
| SEQIDNO_11_Datisca_glomerata | FDLNIPA | 85 | 100 |
| SEQIDNO_13_Glycine_max | FDLNIPA | 85 | 100 |
| SEQIDNO_15_Medicago_sativa | FDLNLPA | 85 | 100 |
| SEQIDNO_17_Nicotiana_tabacum | FDLNMPA | 85 | 100 |
| SEQIDNO_19_Oryza_sativa | FDLNLPA | 85 | 100 |
| SEQIDNO_37_Oryza_sativa | FDLNLPA | 85 | 100 |
| SEQIDNO_23_Triticum_aestivum | FDLNIPA | 85 | 100 |
| SEQIDNO_27_Arabidopsis_thaliana | IDLNLPA | 85 | 100 |
| SEQIDNO_39_Arabidopsis_thaliana | LDLNLTP | 100 | 83,3 |
| SEQIDNO_50_Arabidopsis_thaliana | LDLNLPA | 85 | 100 |
| SEQIDNO_21_Petunia_hybrida | FDLNIPA | 85 | 100 |
| SEQIDNO_25_Capsicum_annuum | FDLNIPA | 85 | 100 |
| SEQIDNO_29_Arabidopsis_thaliana | FDLNIPP | 85 | 100 |
| SEQIDNO_33_Arabidopsis_thaliana | FDLNIIP | 100 | 83,3 |

Table 2. Percentage identity of the EAR motif as found in the polypeptides recited by their sequence database accession number in lines 17 to 30 of page 15 and the paragraph bridging page 15 and 16 of the specification and SEQ ID NO 7.

| | Motif | % Identity to SEQ ID NO: 7 (XDLNXXP) | % Identity to SEQ ID NO: 7 XDLNXP |
|-----------|----------|-----------------------------------------|--------------------------------------|
| AAK01713 | FDLNLPAV | 85 | 100 |
| AF332876 | FDLNLPAV | 85 | 100 |
| AAM67193 | FDLNLPV | 100 | 83,3 |
| NP_199131 | FDLNIIPV | 100 | 83,3 |
| AF250336 | FDLNIPPI | 100 | 100 |
| AY034998 | FDLNIPPI | 100 | 100 |
| AY063006 | FDLNIPPI | 100 | 100 |
| CAA64820 | FDLNIPPI | 100 | 100 |
| NM_102538 | FDLNIPPI | 100 | 100 |
| X95573 | FDLNIPPI | 100 | 100 |
| CAA67228 | FDLNIPPI | 100 | 100 |
| CAA67229 | FDLNIPPI | 100 | 100 |
| X98670 | FDLNIPPI | 100 | 100 |
| X98671 | FDLNIPPI | 100 | 100 |
| BAC43454 | FDLNIPPI | 100 | 100 |
| NP_196054 | FDLNIPPI | 100 | 100 |
| BAA05079 | FDLNIPAL | 85 | 100 |
| BAA05076 | FDLNMLPP | 100 | 83,3 |

| | | | |
|-----------|----------|-----|------|
| BAA05077 | FDLNLPP | 100 | 83,3 |
| BAA05078 | FDLNLPP- | 100 | 100 |
| NP_190562 | FDLNLPAD | 85 | 100 |
| NP_188592 | IDLNLPAL | 85 | 100 |
| BAA19926 | LDLNLPA | 85 | 100 |
| BAA20137 | LDLNLPA | 85 | 100 |
| BAA21919 | IDLNIPAT | 85 | 100 |
| BAC43008 | IDLNLPA | 85 | 100 |
| NP_182037 | IDLNLPA | 85 | 100 |
| CAA67233 | LDLNLPA | 85 | 100 |
| CAA67236 | LDLNLPA | 85 | 100 |
| BAA19111 | LDLNLVPS | 85 | 100 |
| BAA21920 | LDLNLPA | 85 | 100 |
| CAA43111 | LDLNLPA | 85 | 100 |
| BAA19110 | LDLNLTP | 100 | 83,3 |
| BAA21923 | LDLNLTP | 100 | 83,3 |
| BAB67879 | LDLNLTPS | 100 | 83,3 |
| BAA21921 | LDLNLTPR | 100 | 83,3 |
| BAA21922 | LDLNLTPR | 100 | 83,3 |
| CAA67230 | LDLNLTP | 100 | 83,3 |
| CAB67667 | LDLNLTP | 100 | 83,3 |
| BAA19114 | LDLNNLPT | 100 | 83,3 |
| BAA21924 | LDLNLTPY | 100 | 83,3 |
| BAA21925 | FDLNLTPY | 100 | 83,3 |
| BAA21926 | FDLNLTPD | 100 | 83,3 |
| CAA18741 | LDLNPPT | 100 | 100 |
| CAB80245 | LDLNPPT | 100 | 100 |

SEQ ID NO: 8 KRSKRXR

Table 3. Percentage identity of the B-box motif as found in the in the amino acid of the sequence listing of the application as filed and SEQ ID NO 8.

| | Motif | % Identity to SEQ ID NO: 8 (KRSKRXR) |
|--------------------------------|-----------|--------------------------------------|
| IP_070_11_Datisca_glomerata | RKRSKRTR | 100 |
| IP_070_13_Glycine_max | RKRSKRSR | 100 |
| IP_070_15_Medicago_sativa | GKRSKRSR | 100 |
| IP_070_17_Nicotiana_tabacum | GKRSKRPR | 100 |
| IP_070_19_Oryza_sativa | RKRSRRQR | 85 |
| IP_070_37_Oryza_sativa | RKRSRRQR | 85 |
| IP_070_23_Triticum_aestivum | RKRSRRQR | 85 |
| IP_070_27_Arabidopsis_thaliana | RKRSKRQR | 100 |
| IP_070_31_Arabidopsis_thaliana | RKRTKRHR | 85 |
| IP_070_35_Arabidopsis_thaliana | RKRTKRQR | 85 |
| IP_070_21_Petunia_hybrida | GKRSKRQR | 100 |
| IP_070_25_Capsicum_annuum | GKRSKRSR | 100 |
| IP_070_29_Arabidopsis_thaliana | SKGRSKRSR | 100 |

| | | |
|--------------------------------|-----------|-----|
| IP_070_33_Arabidopsis_thaliana | CKKRSKRSR | 100 |
|--------------------------------|-----------|-----|

Table 4. Percentage identity of the B-box motif as found in the polypeptides recited by their sequence database accession number in lines 17 to 30 of page 15 and the paragraph bridging page 15 and 16 of the specification and SEQ ID NO 8.

| | Motif | % Identity to SEQ ID NO: 8 (KRSKRXR) |
|-----------|----------|--------------------------------------|
| AAK01713 | KRSRRQR | 85 |
| AF332876 | KRSRRQR | 85 |
| AAM67193 | KKRSKRSR | 100 |
| NP_199131 | KKRSKRSR | 100 |
| AF250336 | KRSKRSR | 100 |
| AY034998 | KRSKRSR | 100 |
| AY063006 | KRSKRSR | 100 |
| CAA64820 | KRSKRSR | 100 |
| NM_102538 | KRSKRSR | 100 |
| X95573 | KRSKRSR | 100 |
| CAA67228 | KRSKRSR | 100 |
| CAA67229 | KRSKRSR | 100 |
| X98670 | KRSKRSR | 100 |
| X98671 | KRSKRSR | 100 |
| BAC43454 | KRSKRSR | 100 |
| NP_196054 | KRSKRSR | 100 |
| BAA05079 | KRSKRQR | 100 |
| BAA05076 | KRSKRPR | 100 |
| BAA05077 | KRSKRPR | 100 |
| BAA05078 | KRSKRPR | 100 |
| NP_190562 | KRTKRHR | 85 |
| NP_201546 | KRTKRQR | 85 |

SEQ ID NO: 9 EXEXXAXCLXXL

Table 5. Percentage identity of the L-box motif as found in the in the amino acid of the sequence listing of the application as filed and SEQ ID NO 9.

| | Motif | % Identity to SEQ ID NO: 9 (EXEXXAXCLXXL) |
|--------------------------------|--------------|-------------------------------------------|
| IP_070_11_Datisca_glomerata | EEEYLAFCLIML | 100 |
| IP_070_13_Glycine_max | EEEYLALCLIML | 100 |
| IP_070_15_Medicago_sativa | EEEYLALCLIML | 100 |
| IP_070_17_Nicotiana_tabacum | EEEYLALCLIML | 100 |
| IP_070_19_Oryza_sativa | EEENLALCLLML | 100 |
| IP_070_37_Oryza_sativa | EEENLALCLLML | 100 |
| IP_070_23_Triticum_aestivum | EEENLALCLLML | 100 |
| IP_070_27_Arabidopsis_thaliana | EEEYLALCLLML | 100 |
| IP_070_31_Arabidopsis_thaliana | EEEYLALCLLML | 100 |
| IP_070_35_Arabidopsis_thaliana | EEEYLALCLLML | 100 |

| | | |
|--------------------------------|--------------|------|
| IP_070_50_Arabidopsis_thaliana | EEEDMAICLIML | 100 |
| IP_070_46_Arabidopsis_thaliana | VEDTAAKCLMLL | 83,3 |
| IP_070_21_Petunia_hybrida | EEEYLALCLIML | 100 |
| IP_070_25_Capsicum_annuum | EEEYLALCLIML | 100 |
| IP_070_29_Arabidopsis_thaliana | EDEYIALCLMLL | 100 |
| IP_070_33_Arabidopsis_thaliana | EEEYLAFCLMLL | 100 |

Table 6. Percentage identity of the L-box motif as found in the polypeptides recited by their sequence database accession number in lines 17 to 30 of page 15 and the paragraph bridging page 15 and 16 of the specification and SEQ ID NO 9.

| | Motif | % Identity to SEQ ID NO: 9 (EXEXXAXCLXXL) |
|-----------|--------------|-------------------------------------------|
| AAK01713 | EEENLALCLLML | 100 |
| AF332876 | EEENLALCLLML | 100 |
| AAM67193 | EEEYLAFCLMLL | 100 |
| NP_199131 | EEEYLAFCLMLL | 100 |
| AF250336 | EEEYLAFCLMLL | 100 |
| AY034998 | EEEYLAFCLMLL | 100 |
| AY063006 | EEEYLAFCLMLL | 100 |
| CAA64820 | EEEYLAFCLMLL | 100 |
| NM_102538 | EEEYLAFCLMLL | 100 |
| X95573 | EEEYLAFCLMLL | 100 |
| CAA67228 | EEEYLAFWLMLL | 100 |
| CAA67229 | EEEYLAFWLMLL | 100 |
| X98670 | EEEYLAFWLMLL | 100 |
| X98671 | EEEYLAFWLMLL | 100 |
| BAC43454 | EDEYIALCLMLL | 100 |
| NP_196054 | EDEYIALCLMLL | 100 |
| BAA05079 | EEEYLALCLIML | 100 |
| NP_190562 | EEEYLALCLLML | 100 |
| NP_201546 | EEEYLALCLLML | 100 |
| NP_188592 | EEEYLALCLLML | 100 |
| BAA19111 | DEEETARYLILM | 75 |
| BAA21920 | EEEDMANCLILL | 100 |
| CAA43111 | EDEDMANCLILL | 100 |
| BAA21921 | EAAAMANCLMLL | 91,6 |
| BAA21922 | EAEAMANCLMLL | 100 |
| CAA67231 | VDVTAANCLMLL | 83,3 |
| CAA67232 | VDVTAANCLMLL | 83,3 |
| CAA67234 | EEDTTAKCLMLL | 91,6 |
| CAA67235 | EEDTTAKCLMLL | 91,6 |
| CAB90936 | EEDTTAKCLMLL | 91,6 |
| CAB90935 | VEDTAAKCLMLL | 83,3 |
| BAA21925 | ETLAMANCVNIL | 83,3 |
| BAA21926 | ERLAMENCANIL | 83,3 |